Ultrafast Protein Dynamics Revealed by Infrared Spectroscopy

van Wilderen, L.J.G.W.

2007

document version
Publisher's PDF, also known as Version of record

Ultrafast Protein Dynamics Revealed by Infrared Spectroscopy

Luuk Joost Gerard Wouter van Wilderen

VRIJE UNIVERSITEIT

Ultrafast Protein Dynamics Revealed by Infrared Spectroscopy

ACADEMISCH PROEFSCHRIFT

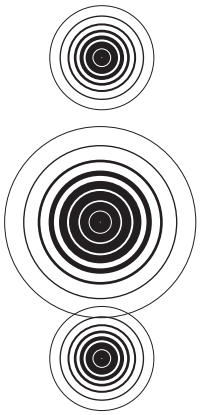
ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Exacte Wetenschappen
op woensdag 27 juni 2007 om 10.45 uur
in de aula van de universiteit,
De Boelelaan 1105

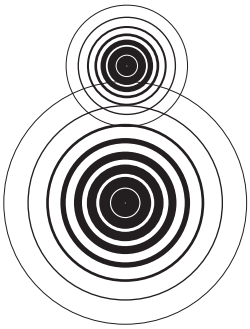
door

Luuk Joost Gerard Wouter van Wilderen

geboren te Tegelen

promotoren: prof.dr. R. van Grondelle
prof.dr. M.L Groot





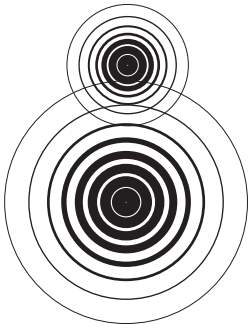
Ultrafast Protein Dynamics Revealed by Infrared Spectroscopy

Luuk Joost Gerard Wouter van Wilderen

The research described in this thesis was performed in the Department of Biophysics of the Faculty of Exact Sciences, Vrije Universiteit, De Boelelaan 1081, 1081 HV, Amsterdam, The Netherlands.
ISBN 978-90-8659-104-6

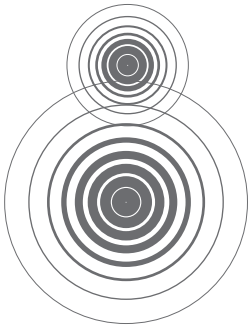
This research was financially supported by NWO-ALW, the Dutch Organization for Scientific Research via the Dutch Foundation for Earth and Life Sciences.

This thesis was reviewed and approved by:
Prof. dr. R.M. Hochstrasser, University of Pennsylvania, USA
Dr. E.T.J. Nibbering, Max Born Institute, Germany
Dr. J.J. van Thor, University of Oxford, UK
Dr. J.T.M. Kennis, Vrije Universiteit Amsterdam
Prof. dr. K.J. Hellingwerf, Universiteit van Amsterdam
Printed by: Gildeprint Drukkerijen B.V., Enschede
Layout and Graphic Design: Daniela Bellelli



to my family

Contents



Introduction

01

Abstract

This chapter serves to provide an introduction to the field of ultrafast infrared spectroscopy and proteins. In addition, background and motivations of this study are highlighted. The subjects of study, GFP, PYP and FixL, are three light sensitive proteins which are briefly introduced. Finally, the fundamental principles of the experimental technique and methods of data analysis are described.

Introduction

Earth is crawling with life. Communication and interaction between individuals and groups of organisms can be highly complex and viewed from different levels. Very diverse and complicated systems can be adequately described by relatively simple rules and laws (1, 2). On the level of a single (micro)organism, even the smallest parts (such as proteins or even smaller) of the organism play a role in an incredible network of interactions. These also can be described with a set of rules, but on a different level. The interactions discussed here can be described by the laws of physics, mechanics and thermodynamics at the cellular level and electrostatics and quantum mechanics within proteins. A basic 'low–level' understanding of protein function is needed to be able to solve or prevent 'upper–level' problems such as metabolic diseases (such as thyroid diseases). Therefore, protein model systems must be used to unravel the underlying mechanisms step–by–step. Still, it is a formidable task to decipher how proteins sense changes in their environment and ultimately generate interpretable and biologically relevant signals (to, for example, a cell or a bacterium). Nowadays it is possible with advanced molecular dynamics simulations to predict time–dependent protein dynamics on short time–scales (3, 4). However, the results of these calculations depend on starting assumptions and the quality of the starting structure fed into the program, which must be determined experimentally, may be unavailable, incomplete, or even wrong. Hence, experiments are always needed to check the validity and can in turn be used to fine–tune the simulations and even help to interpret them. Theory and experiment mutually interact and contribute to a full understanding of the actual physical world. In this thesis some of the basic rules that govern protein function within an organism are discovered and described. First a short introduction to proteins in general is given, and then the model proteins GFP, PYP and FixL are described in detail. The experimental and computational approaches to understand these proteins are discussed, and a short introduction to our advanced data –analysis techniques is given.

Proteins

Proteins are like little organic motors. They actively perform all kind of tasks in nature and the human body. They are made up of a long string of amino acids which are folded in a very particular way contingent on their amino acid sequence (5). Misfolded proteins may form a protein without any biological function, or can even be harmful (like prions). It is the interaction between the different parts of the amino acid chain that determine the exact way a protein is folded. Hydrogen bonds between different protein parts are the main interactions responsible for the conformation.

For some proteins, the outside environment has an influence. Some proteins are able to 'sense' the oxygen level in the cell where it resides. Others are sensitive to the acidity, or even the light conditions. In fact, they are highly specific: for instance some light–sensitive proteins cannot only sense light, but also the intensity and colour. Think for example about your eyes, where the rods and cones in your retina are packed with light sensitive proteins. Some of these proteins are sensitive to red light, others to blue or green, again others only to the intensity of light. In nature proteins act as the mediators between the environment and the cell. Most sensory proteins are able to repetitively perform the same task over and over again upon a certain specific input, and to send an interpretable signal to the cell or organism as an output. In this thesis we study three different light–sensitive (photoactive) proteins, which can be used as model systems for protein function in general: Green Fluorescent Protein (GFP), Photoactive Yellow Protein (PYP) and FixL. A light flash induces a process in these proteins on an extremely fast time scale: faster than a millionth of a millionth of a second (a picosecond). Very short laser pulses are therefore needed to study the very first steps in signal generation in a protein, because the shorter the pulses we use, the faster processes we are able to follow.

Green Fluorescent Protein: From Proton Transfer to Bioluminescence

One of the model proteins investigated here is green fluorescent protein (GFP). GFP is found in multiple organisms, but the protein found in the jellyfish *Aequorea victoria* has become a popular fluorescent marker. It is used for imaging in biology and medicine, and thanks part of its popularity to its stability and excellent fluorescent properties (6, 7). The process responsible for this green fluorescence is an ultrafast proton transfer mechanism. It can be fused to another (colour–less) protein so that the whereabouts of the other protein can be monitored by the bright green fluorescence of GFP. Inside the jellyfish, together with the protein aequorin, GFP is responsible for its bioluminescence. Bioluminescence in organisms usually has an offensive, defensive or communication function. Why this species of jellyfish uses GFP is however not known. It has been suggested to be a non–membrane bound proton pump; by attaching itself to a membrane it transfers protons from one side to the other side (hence acidifying it) (8). Aequorin, the partner of GFP inside the jellyfish, emits blue light upon binding of Calcium ions. This blue light is consequently absorbed by GFP, which, in turn, emits green light. In this work, only GFP is studied. The crystal structure of GFP (9) provides some clues to the fluorescence mechanism of GFP. Wild type GFP (the naturally occurring protein) has revealed a unique barrel–shape comprised

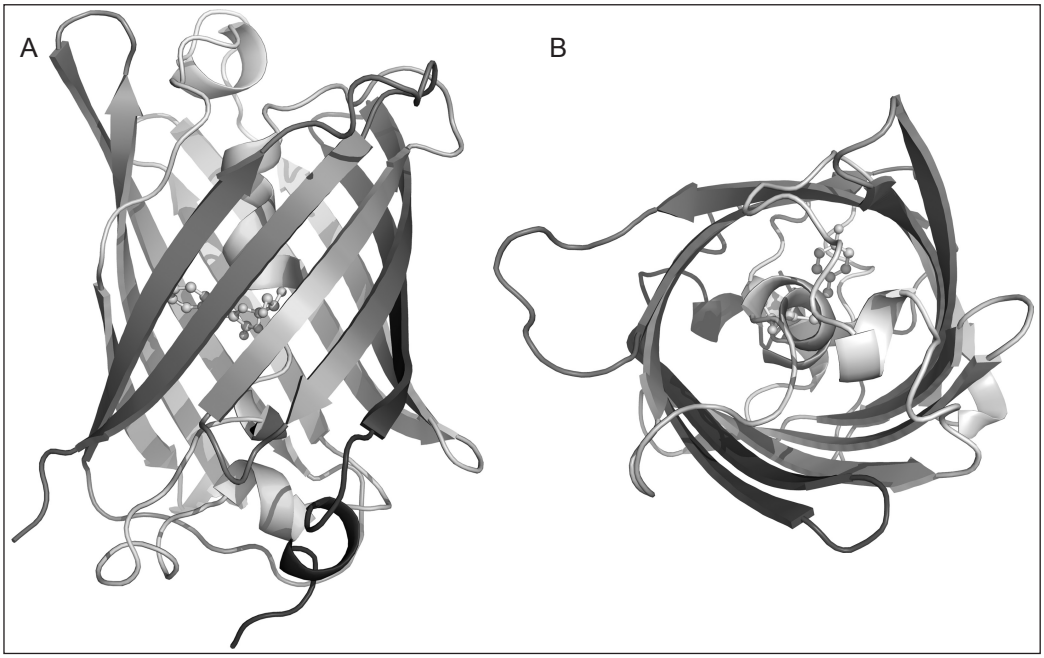
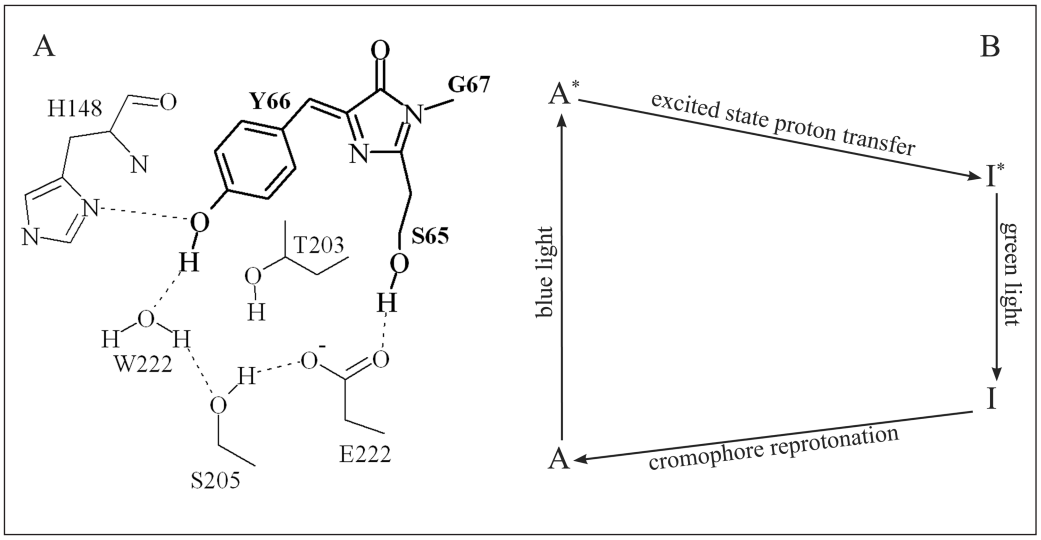


Figure 1 **Schematic representation of GFP; a side (left) and a top–view (right). The protein is drawn in ribbons, the buried chromophore with balls–and–sticks.**

of 11 beta–sheets containing one internal alpha–helix (Figure 1). The light–absorbing part of GFP is a pigment molecule (the chromophore, depicted as a ball–and–stick molecule in Figure 1), which lies protected inside this barrel. The chromophore is shown to be 4–(*p*–hydroxybenzylidene) imidazolid–5–one (10), and is schematically drawn in Figure 2A.

The GFP chromophore is unique in that it is formed auto–catalytically from the three amino acid residues Ser65–Tyr66–Gly67 (11). Upon the absorption of light GFP follows a photocycle with distinct intermediate steps, each with its own spectroscopic properties. Upon blue–light absorption (the protein absorbs in the near–UV at ~400 nm), the chromophore becomes electronically excited. Subsequently, it fluoresces with a high quantum yield (i.e. photocycle efficiency) of ~0.8 in the green part (around 510 nm) of the spectrum (see Figure 2B). When the cycle is complete (in less than 4 nanoseconds), the protein has returned to its original configuration and can be reused again by the jellyfish, see Figure 2B.

The photocycle of GFP, depicted in Figure 2B, is actually surprisingly complex. For example, it has been shown by visible pump–dump–probe experiments that the intermediate I actually consists of I_1 and I_2 (12). The first I_1 state decays so fast that with 'normal' pump–probe experiments it remains virtually invisible. Besides the A– and the I–forms, another protein state is produced with a low fraction, referred to as the B–form (13). The B–form is also formed upon excitation of A, but this state does not return to the A–state; it has a long, excitation wavelength–dependent



lifetime on the order of days (14, 15). It has been shown that the B–form is actually a photo–damaged protein. The glutamic acid residue E222 is decarboxylated and has split off a CO_2–molecule (13). Therefore, once the B–state is formed, this protein fraction is permanently lost.

Strangely, the chromophore in solution is scarcely fluorescent (11, 16), and the formation of a twisted conformation has been hypothesized to quench the fluorescence (17, 18). The protein environment has therefore a profound effect on the chromophore dynamics. Inside the protein, it is the deprotonation of the excited–state chromophore (the phenolic moiety looses its proton via a 'proton–wire') that is believed to cause the large spectral shift (19). The magnitude of this spectral shift is exactly why GFP is widely used as a fluorescent tag.

The mechanism for the proton transfer process is generally thought to start with the proton at the phenol–moiety of the chromophore, and ends with the proton arriving at the glutamic acid E222 residue (see Figure 2A). The proton is thought to 'hop' over a proton–wire, via a water molecule and Ser205 to the terminal proton acceptor Glu222, and then it shuttles back to the chromophore (20, 21). Model calculations have however proposed different proton pathways. The proton transfer event may start from the chromophore (22) or start from the end, i.e. from Glu222 (23). Recent electronic structure calculations by Vendrell *et al.* found that a pulling of the proton over the wire initiated by transfer of a proton from the Ser205 to Glu222 is energetically most favourable (23). A concerted mechanism has been found by other recent quantum mechanical modelling (24, 25). Infrared pump–probe measurements (chapter 2) are conducted to reveal the nature of proton transfer in native GFP.

Mutants of GFP have been created which have changes

Figure 2 **Schematic drawing of the chromophore–binding pocket of GFP in panel A (state 'A' in panel B, adapted from ref. Jung *et al.* (59), based on Brejc *et al.* (21)). The chromophore (thick lines) is build up out of three amino acid residues (Y66–G67–S65) and is embedded in a hydrogen–bonding network. In panel B a schematic energy diagram is shown for the GFP photocycle (adapted from (12)). Upon blue light excitation the ground state A is electronically excited and forms the excited I*–state by deprotonation of the chromophore. Upon decaying to I the green fluorescence is emitted.**

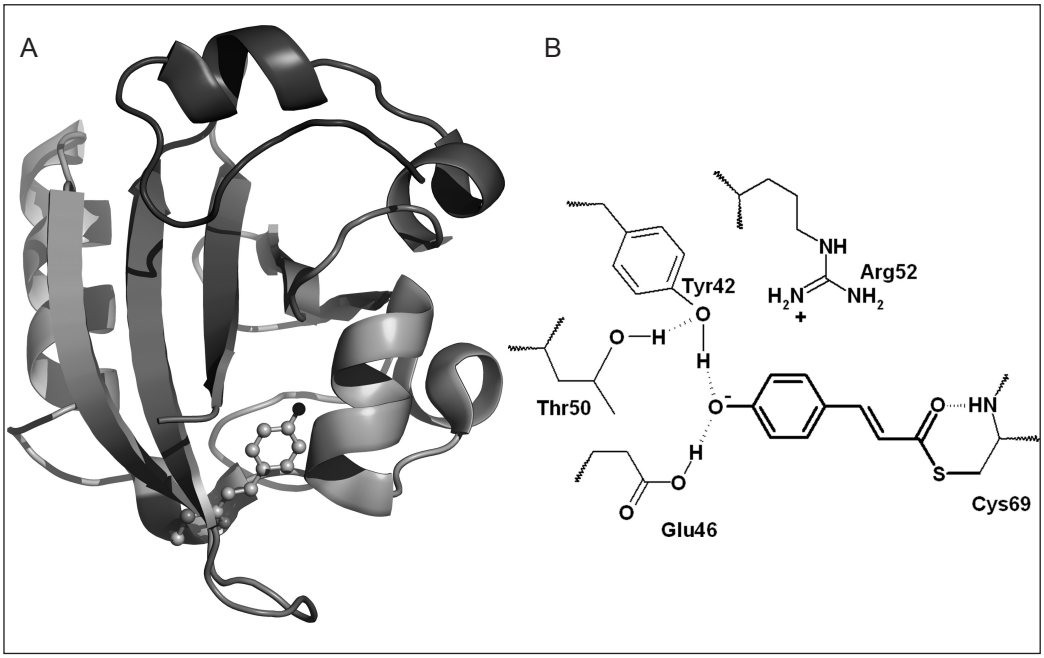


Figure 3 **Model representation of PYP with the ribbon–model in panel A. The chromophore is drawn with the ball–and–stick model in A and schematically drawn in B (thick lines) with the directly surrounding amino acids. The chromophore is covalently bound to protein backbone by means of Cys69, and embedded in a hydrogen–bonding network with Glu46 and Tyr42 (also hydrogen bonded to Thr50). The positively charged Arg52 residue possibly stabilizes the negative charge on the chromophore.**

in key residues and produce a range of available fluorescent colours (26, 27). Methods to prolong and tweak the fluorescence characteristics are continuously investigated, and a fundamental understanding of the protein's photocycle pathway is therefore necessary. These experiments should shed light on not only the photocycle mechanism of GFP, but also on protein function in general, as proton–wires have important biological functions ranging from pumping protons over membranes (28–31), to the catalytic function of enzymes (32, 33).

Photoactive Yellow Protein: From Isomerization to Phototaxis

Photoactive Yellow Protein (PYP) is the second model system investigated in this work (PYP, see Figure 3A). PYP is a small protein (125kDa, 125 amino acids) that is found in the swimming halophilic bacterium *Halorhodospira halophila* (34). The bacterium requires light to perform photosynthesis and to survive, but too much (blue) light is harmful. The bacterium is negatively phototactic, which means that it is able to sense the intensity of light and, according to its needs, swim toward or away from it (35). The behaviour of the bacterium matches the spectral properties of PYP, and therefore it is thought to be the mediator between the environment and an interpretable signal for the bacterium. PYP is a blue–light receptor, which means it can absorb blue light. Upon the absorption of light PYP follows a photocycle (Figure 4), just as GFP. The chromophore, a *p*–coumaric acid (the thick lines in

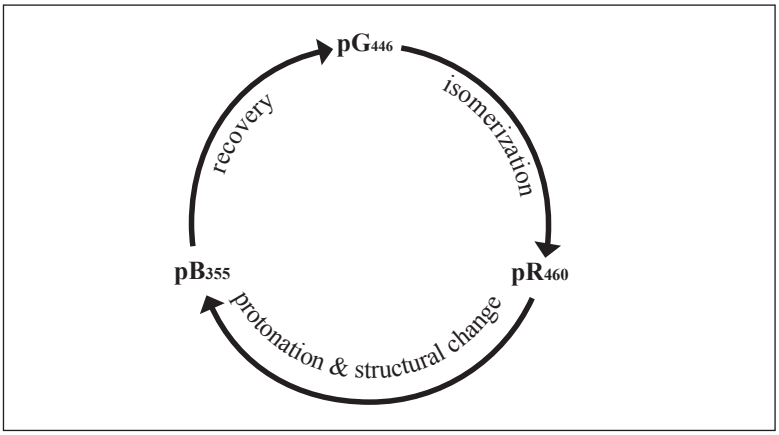


Figure 4 **Simplistic model for the photocycle of PYP, where pG denotes the ground state, pR is the red–shifted state and pB the blue–shifted intermediate (adapted from (34)). The numbers represent the main absorption peak in nanometres of the individual species.**

Figure 3B), isomerises about its central double bond, changing the chromophore conformation from *trans* to *cis*. The isomerization consequently leads to conformational changes in the protein as a whole (forming the 'signalling state'), and ultimately the protein returns to its original conformation (the ground state).

The photocycle of PYP has been investigated using many experimental techniques which have given a detailed insight into its mechanism (34, 36, 37). In addition, model calculations and model chromophore studies in solution have resulted in a nearly complete picture. Wild type PYP has a photocycle that spans more than 14 decades in time from picoseconds to milliseconds depending on temperature, acidity, and humidity (34). The cyclical nature of the photochemistry allows the bacterium to reuse the protein over–and–over again (34), see Figure 4. In a similar manner to GFP, the protein environment has a profound effect on the chromophore dynamics, as the differences resulting from ultrafast (femto– to picosecond) dynamics of photo–excited model chromophores in solution illustrate. The exact conformation of the chromophore model and the effect of conformation on photoisomerization has been the subject of considerable debate (38, 39). In wild type PYP, the chromophore consists of a negatively charged *p*–coumaric acid (hydroxy–cynnamic acid) (40). Inside the chromophore–binding pocket the available conformations of the chromophore are limited by the surrounding residues and their hydrogen–bonding network. The residues and their interactions also tune the absorption properties of the protein (41), and give it its characteristic yellow colour (absorbing mainly at 446 nm).

Upon excitation, the negative charge on the phenolic ring of the chromophore moves towards the carbonyl (42). The conjugation across the chromophore is subsequently rearranged, leading to *trans–cis* isomerization about the C–C double bond. The excited state lifetime of PYP has been determined by fluorescence spectroscopy and found to be multi–exponential and on the order of picoseconds (43, 44). This is also the time scale (1–2 ps) on which,

with time–resolved vibrational spectroscopy, a stable *cis*–ground state has been observed to form (45, 46). In conjunction, the visible absorption spectrum is shifted to the red. The photocycle continues by a protonation of the phenolic moiety of the chromophore and a structural rearrangement of the protein (absorbing at ~355 nm) (34, 36, 37). This is the proposed signalling state, i.e. the biologically productive state. In this state it may interact with a downstream signalling partner, but as or yet, this component has not been identified. Consequently, the chromophore deprotonates, reisomerizes, and the protein structurally rearranges to form the original configuration. The photocycle (in Figure 4) is, like GFP, surprisingly complex. Up to eight intermediates have been determined up to date (34). In this study we have focused on the initial part of the photocycle spanning the time range from 150 fs to 3 ns, up to the signalling state. The results of our investigations are described in chapter 3 and 4. The first 3 nanoseconds of the photocycle are key to the formation of the signalling state, and will be treated in chapter 3. Advanced spectroscopic techniques such as pump–dump–probe experiments (the excited state population is optically de–excited to the ground state, (47)) in the visible part of the spectrum have identified several ground state intermediates. These states are formed with a relatively high quantum yield (~50 %, (47)), while the biologically productive state is only formed with ~30 % (45, 47, 48) (the remaining ~20% thermally decays back into the ground state). Potential structural interactions that play a role in this photocycle discrimination process are the hydrogen bond between the chromophore's phenol oxygen and Glu46, and the hydrogen bond with the chromophore's carbonyl. Both interactions are examined by the use of an E46Q PYP mutant, and can be excellently studied by infrared spectroscopy (see chapter 4**)**. A detailed understanding of the chemical kinetic mechanism of a protein (along with a chemical basis for its efficiency) is important for the understanding of protein function as they related to health and disease.

FixL: An Oxygen Sensor

The protein FixL responds to fluctuating oxygen levels in the bacterial environment, and is found in the nitrogen–fixing bacterium *Bradyrhizobium japonicum* (and therefore referred to as *Bj*FixL) (49). *Bj*FixL is composed of two domains (a compact, folded region of an amino acid chain that may be connected to another one by a flexible segment of polypeptide chain (5)), an N–terminal heme–PAS sensory domain and a C–terminal signaling kinase domain. In this thesis we focus on the heme–PAS–domain, referred to hereafter as *Bj*FixLH. Structurally, *Bj*FixLH is similar to PYP, another member of the PAS–domain family. However, instead of a *p*–coumaric acid FixL has a heme in the chromophore–binding

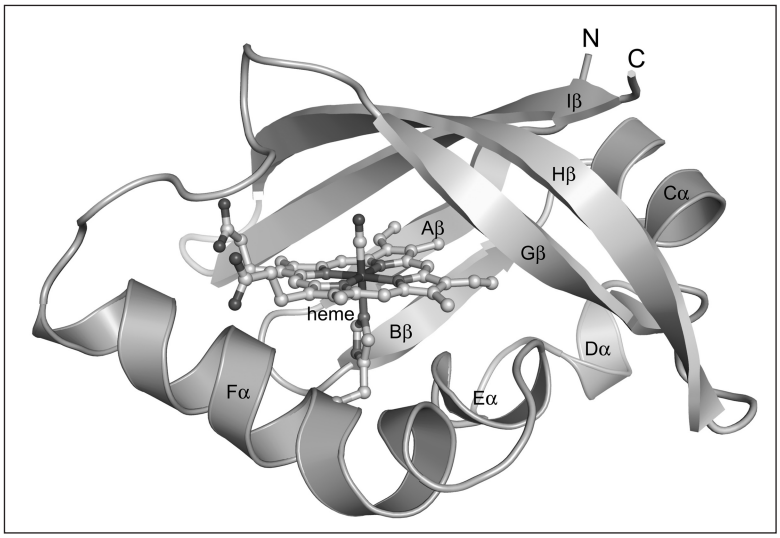


Figure 5 **Schematic representation of FixL drawn in ribbons (created by J. Key). The heme–group is drawn in balls–and–sticks, with the CO ligand (also in balls–and–sticks) sticking out vertically to the top of the centre of the heme.**

site (see Figure 5). Heme–PAS proteins have been identified in a range of organisms, ranging from bacteria (50) to mammals (51) (see (52) for a review). The heme group can bind a diatomic molecule (the ligand; for example CO, NO or CN^-) to its central iron atom, and is therefore similar to myoglobin and hemoglobin, the oxygen transporters for the cells in the human body.

*Bj*FixLH In the deoxy state (Fe^{2+}; without ligand) has a five–coordinated heme group that is buried in a hydrophobic pocket. Ligand binding results in a six–coordinated configuration of the heme iron atom. Crystallographic studies have revealed that, upon binding of oxygen to deoxy *Bj*FixLH, a conformational change occurs in the FG loop and in the H and I strands of the β sheet of the protein (53). The side chain of Arg220 moves into the heme pocket where it serves as a distal coordinating residue to the oxygen ligand. Carbon monoxide, however, does not cause the same change in Arg220, yet structural changes are evident in the β strands of the protein, a region implicated in signal transduction in other PAS domains of diverse function. Up to date, little is known about the dynamics of CO photolysis (the release of CO upon photo–excitation) within the widely distributed heme–PAS domain. Therefore, the first steps in signal transduction of CO–*Bj*FixLH are studied by ultrafast pump–probe spectroscopy in the infrared. The infrared frequency of the CO molecule is very sensitive to interactions with the surrounding protein, and acts therefore as a probe for the chromophore–binding pocket dynamics. The heme–protein myoglobin contains cavities surrounding its heme where the ligand transiently docks upon photolysis of the heme–ligand bond (54). The trajectory from heme to solvent may involve similar binding sites in *Bj*FixLH as well, and time–resolved infrared experiments can provide the necessary structural information to demonstrate any similarity with the ultrafast dynamics in myoglobin. The very

first steps upon ligand dissociation play an important role in the signal transduction process, and into the common physiological process of ligand release. Our experimental results are presented in chapter 5.

Experimental Method: Infrared Pump–Probe Spectroscopy

To investigate the fundamental biological process in Nature described above, isomerization, proton transfer, and ligand dissociation, special experimental techniques have to be employed. These processes occur on incredibly fast time–scales and ultrafast lasers are necessary to observe them. In its natural environment proteins are 'waiting' for a physiological input or change. Using lasers it is possible to make the protein perform its tricks when desired by the experimentalist.

The main spectroscopic technique used in this thesis is infrared difference absorption or pump–probe spectroscopy. The fundamental principles are depicted in Figure 6. The basis of infrared difference absorption is an accurate measurement of the change in the optical absorption, or optical density OD, of the sample over time. The absorbance at a certain wavelength λ of a sample containing pigments (colour absorbing molecules) is described by

$$OD(\lambda) = -\log \frac{I(\lambda)}{I_0(\lambda)}$$

where I_0 is the incident light intensity and I the transmitted light intensity.

The photoreaction in the sample is initiated with a laser pulse of the appropriate energy (the pump), and subsequent changes in the optical properties of the sample are measured by a probe beam. It is possible to relate these changes to dynamics of the protein. The probe beam contains a spectrum as wide as experimentally attainable to monitor the spectral changes in the greatest possible wavelength range, resulting in more information. The probe has a relatively low intensity with respect to the pump–beam so that it does not start any photoreaction in the sample. It is possible to monitor the absorption changes in the sample as a function of time by using a delay line, which consists of two mirrors on a mechanically moveable stage. Light travels with a speed of $3 \cdot 10^8$ m/s in vacuum; creating a longer path length for the (laser–)light to travel increases the time it takes to reach the sample. The smaller steps (in distance) these mirrors can make, the higher our time–resolution. However, the final time–resolution is actually limited by the cross–correlation of the pump and probe beam and not by the step size of the movable stage (resulting in a time–resolution

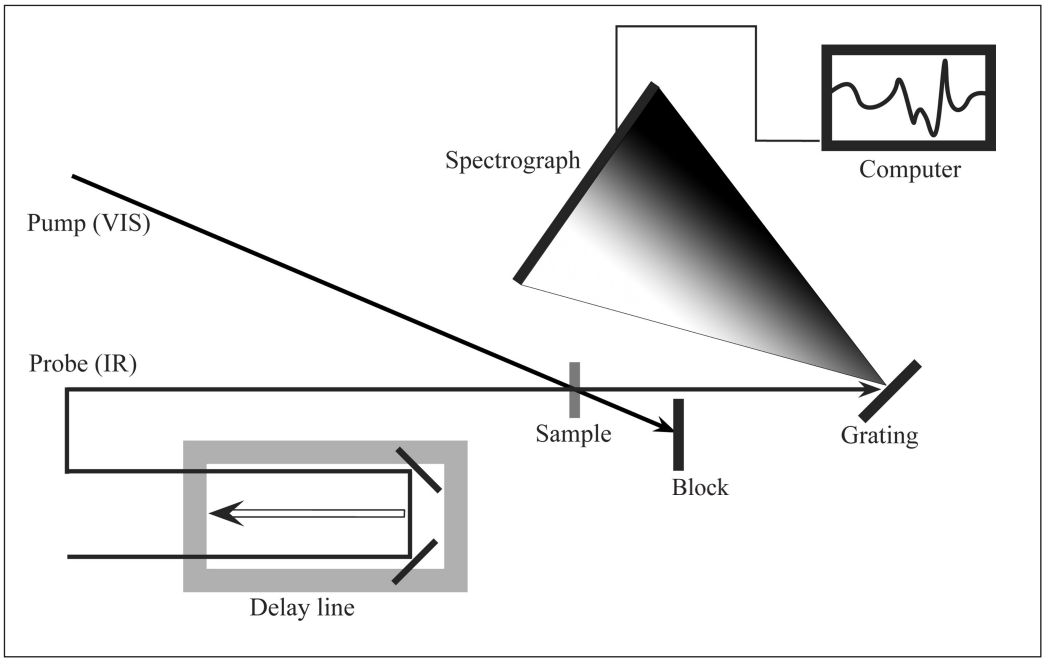


Figure 6 **Schematic drawing of a pump–probe experiment.**

of ~150 femtoseconds). After reaching the sample, the pump–beam is blocked while the probe is dispersed through a grating onto a detector and stored by a computer. The time–dependent absorbance spectra for the 'pumped' sample is compared with the 'un–pumped' sample. In this manner we can measure the pump–induced infrared absorption difference as a function of time.

The experimental setup is shown in detail in Figure 7. A Ti: Sapphire oscillator–regenerative amplifier laser system (Hurricane, Spectra Physics Inc.) produces 800 nm light at a 1 kHz repetition rate. The pulses are 80–90 fs in length and have an energy of about 0.8 mJ. Part of the light is diverted with a beamsplitter to pump a non–collinear optical parametric amplifier (NOPA), which can be tuned to produce 60 fs excitation pulses with a wavelength range of 475–800 nm. The polarization of the excitation pulse with respect to the infrared probe pulse is set with a Berek polarizer. The remaining 800 nm light pumps a tuneable optical parametric generator and amplifier with difference frequency generator (TOPAS, Light Conversion). Spectrally broad mid–infrared pulses (width ~200 cm^{-1}) are produced with a central wavelength between 2.4–11 μm, having an energy of 3–15 μJ. The beam is attenuated (symbolized by the OD 2 filter in Figure 7) to about 1 nJ before it reaches the sample. The probe has a diameter of about 120–150 μm, the pump is the same or slightly larger. After the sample the probe is dispersed on a spectrograph (Chromex) and detected by a 32–element Mercury Cadmium Tellurium detector (MCT, made by Infrared Associates). Three gratings with 150, 75, or 50 grooves per mm can be used with this spectrograph

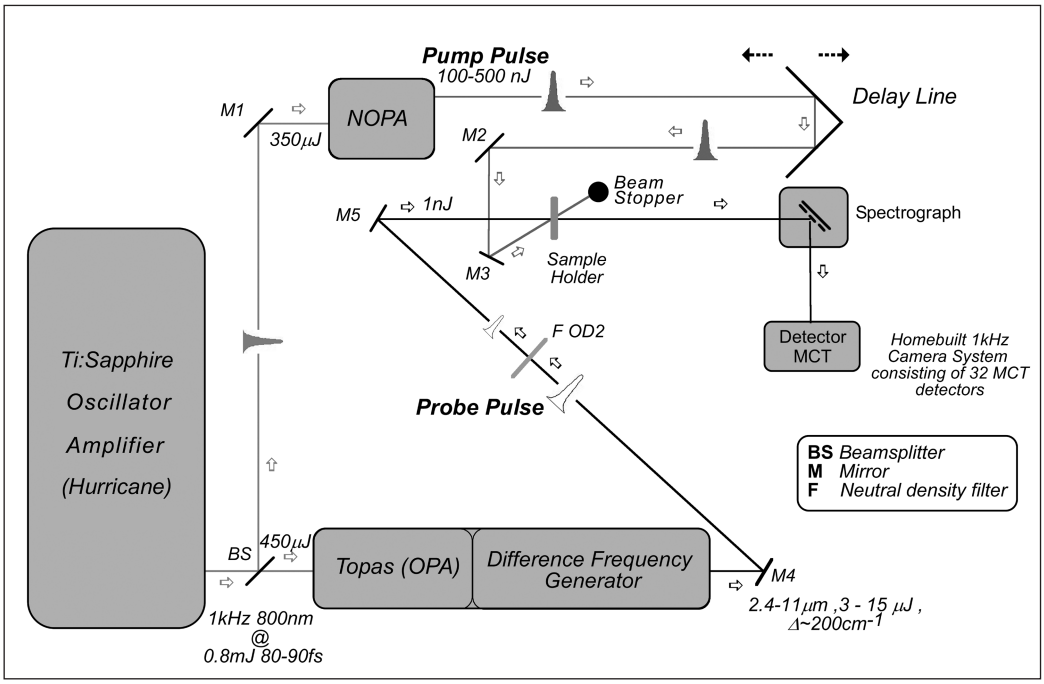


Figure 7 **Schematic layout of the femtosecond visible pump–mid–infrared probe setup (created by C. Bonetti). The arrows denote the direction of travel of the beam. In addition, some laser pulse characteristics are shown.**

to obtain a spectral resolution of ~6 cm^{-1} or more or less. All 32 channels are individually amplified, and read in by 32 home–built integrate–and–hold (IH) devices. A pc with a data–acquisition card (PCI6031E, National Instruments) is used to collect and record the signals from the IH–box. The infrared probe path is entirely placed inside a nitrogen gas or dry air purged box to reduce absorption due to water vapour. A phase–locked chopper (500 Hz) blocks every other pump–laser shot, so that the transmission with and without pump can be recorded. In such fashion the pump–induced change in transmission can be determined and the difference absorption can be calculated.

Spatial and temporal overlap of the visible and mid–IR beam is obtained using a 50 μm thick slab of GaAs. The visible excitation pulse creates free electron carriers in this material which results in a large change in the index of refraction. Thus huge changes in the transmission of the IR pulse (corresponding to several hundreds of mOD of absorption change) are visible. The cross correlation of the visible and IR pulses as measured in GaAs is typically about 180 fs. In a single experiment, a spectral probe window of about 180 cm^{-1} (when a spectral resolution of 6 cm^{-1} is sufficient) is obtained. Partly overlapping spectral regions must be measured to obtain a wide spectral window.

Infrared spectroscopy is very sensitive to structural changes such as isomerization or changes in hydrogen bonding. It is possible to follow a change in vibrational frequency of one single bond as its interaction with its environment changes. It is the

perfect tool to study biological systems such as proteins. Consider for example two molecules that interact with each other by forming a hydrogen bond between them. If this hydrogen bond is broken, the frequencies of the two interacting bonds then up–shift to higher frequencies (compared to the original ground–state configuration). It can be understood by viewing an atomic bond as a spring, and the natural frequency it vibrates in is characterized by the spring constant. Less interaction with the other molecule, breaking the hydrogen bond between them, causes the atomic bonds to become shorter and vibrate faster (increasing the frequency).

With the experimental setup described above, it is also possible to determine the time–dependent orientation (the anisotropy) of individual atomic groups in a protein. Most pump–probe measurements performed in this thesis are performed under magic angle conditions, meaning that the pump–beam has a polarization rotated over 54.7^{o} with respect to the probe beam, and resulting in anisotropy–free measurements (55). In chapter 6 measurements on PYP are described where the polarization of the pump–beam with respect to the probe beam is changed to perpendicular, magic angle and parallel, to probe the change in orientation from individual atomic groups relative to the optical ground state transition dipole moment. Of particular importance are the isomerization–dependent orientations of the chromophore's carbonyl group and of the Glu46 amino acid residue. The results of these experiments are done to support the experimental data presented in chapters 3 and 4.

Calculations

Infrared spectra are packed with information about the vibrational structure of the chromophore and the protein. To determine the exact identity of one specific band, a simple measurement of the spectrum is not sufficient. Time–resolved data reveal information about the evolution in time of a band and is highly informative, but only gives a hint about its true origin. In general, large–scale protein motions (such as for instance the translocation of a complete helix) do not occur on very rapid time–scales and so are not responsible for bands evolving in less then 1 nanosecond. The players in the sub–nanosecond domain are therefore the chromophore and the neighbouring amino acid residues in the chromophore–binding pocket. But also chromophores can have such crowded spectra that its vibrations practically dominate the whole mid–Infrared spectrum (such as is the case for GFP, chapter 2). A time–consuming and costly way is to isotope–label one atom in the protein and see what effect it has on the spectra. Or one residue can be mutated into another residue, or just the chromophore in solution can be analyzed. Another approach is to use available protein structures to calculate a normal mode analysis

and to calculate the vibrational spectrum of the chromophore. It gives the exact nature of one vibration, although the intensity and location are still not 100% reliable (56). However, calculations do give reliable results trend–wise (one vibration is for example always higher in frequency than another one), and are therefore an extra tool to help to understand the experimental data.

Different computational methods are commonly used to perform vibrational calculations, such as density functional theory (DFT) and *ab initio* (based on first principles) Hartree–Fock (HF) methods (56, 57). HF uses the assumption that each electron sees all of the others as an average field, thus neglecting electron correlation and interaction. DFT in combination with an electron correlation functional (such as B3LYP) includes these interactions and results generally in better results, but renders the calculation more time–consuming (56).

Currently, HF and DFT are between the most cost–effective methods: relatively fast calculations with high accuracy. However, a scaling factor is still necessary due to a standard overestimation of the frequencies, although it can approach unity (57). The main reasons are the neglect of anharmonicity, and the amount of basis–sets chosen (the amount of mathematical functions used to describe the atoms of the studied molecule and the interactions between the atoms).

Ideally, these calculations would be done on the whole protein, but this is currently impossible due to the lack of computer resources. Up to approximately 25–30 atoms in total is however feasible with a fast and up–to–date desktop computer, but including all interacting (such as hydrogen bonds) and spherically hindering residues in the chromophore–binding pocket will still virtually take forever. An often–used trick to obtain more biologically relevant results while reducing the overall computational effort is, for instance, to mimic the hydrogen–bonding interactions between the amino–acid residues in the chromophore–binding pocket by (relatively small) water molecules.

Before a vibrational (normal–mode) analysis is done, a geometry optimisation has to be done. The reason is to find a minimum on the potential energy surface (PES). This is necessary because the used coordinates (usually coming from X–ray structures) are frozen structures, but real molecules thermally vibrate. Therefore, X–ray structures represent most probably not a minimum on the PES. Also, hydrogen atoms are not visible in these structures and have to be manually inserted. Performing a normal–mode analysis on a non–geometry optimised structure (or with a different level of theory) does not have any chemical meaning and must therefore be checked for the absence of imaginary frequencies. If imaginary frequencies are found, the structure the calculation is performed on does not represent a minimum on the PES, and can be a transition–state or a stationary point, i.e. a molecular

geometry that 'feels' a force field and is on its way to change its conformation to another one. The imaginary frequency then represents a coordinate of motion that leads to a lower energy. The calculations performed in this work are done on a model PYP chromophore to study its isomerization process (see chapter 7). The interactions with some residues in the chromophore–binding pocket are mimicked by water molecules. The calculations are done to support the interpretations of the experiments described in chapters 3, 4 and 6.

Data Analysis

An infrared absorption difference experiment results in 3–dimensional data; Changes in absorption intensity are measured as a function of time and wavelength. A typical dataset results in ~20.000 datapoints. Analysing all acquired data simultaneously gives correlations between different wavelength–regions and time scales. This type of analysis is called global analysis (58), and requires some calculational effort. However, in this way it is possible to extract and quantify physical information out of the immense amount of data. A fundamental step is the description of all data with a minimum amount (i) of time constants (k_i) that result in a good–quality fit. Each component starts with a given concentration (c_i) that decays in time and possesses its own specific time–independent spectrum (ε_i). Mathematically, the observed signal S at any given time (t) or wavenumber (λ) can be described with

$$S(\lambda,t)=\sum_{i=1}^{n} c_i(t)\,\varepsilon_i(\lambda)$$

with

$$c_i = e^{(-k_i t)}$$

However, the resulting spectra ε_i do not always contain physically relevant information. Decisions have to be made about a model that not only describes the raw data, but also represents biologically and physiologically relevant information. Previous knowledge of a system such as an already (partially) defined protein photocycle, obtained by other experimental techniques, can assist in the global analysis process. However, always the raw data have to be kept in mind, as one can quickly get lost in different models and interpretations. It is the data that need to speak, and not the desires for an interpretation or explanation of the experimentalist. Therefore the simplest model describing the measured data has to be chosen. The simplest model templates are the sequential and parallel models (see Figure 8A and B), where one compartment

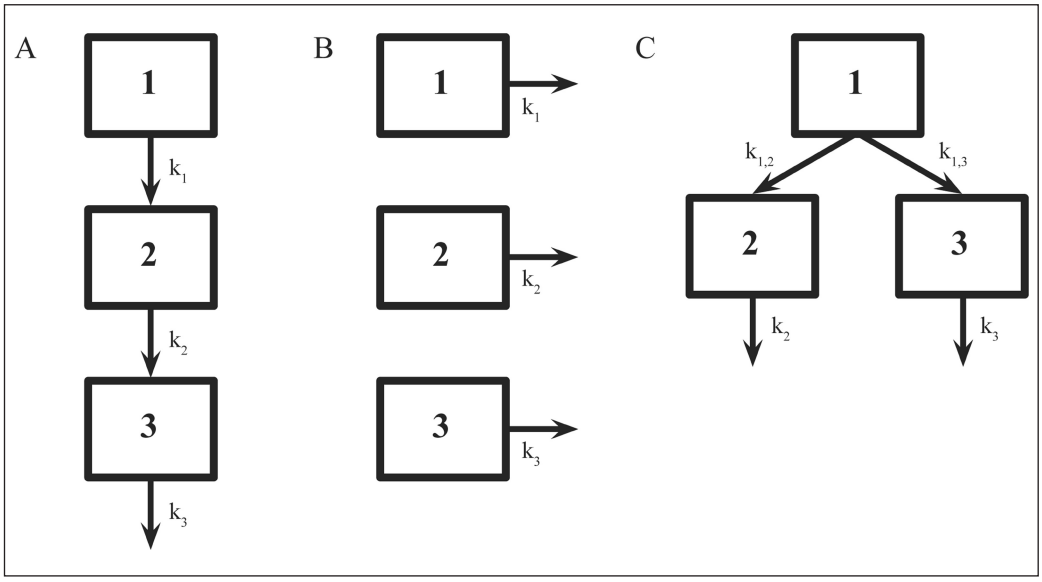


Figure 8 **Schematic view of a sequential model in A, a parallel one in B, and a branched scheme in C to be used with the global analysis method. All models produce different spectral properties for each compartment (except for compartment 1 in A and B) when used to analyse the same data.**

flows directly into the next compartment (or to the ground state in Figure 8B). A compartment represents a spectroscopic distinct state or species. A more complicated model is a branched model (Figure 8C), where one compartment can populate two other compartments.

If one dataset is analysed with the three described models, then the resulting spectra for each compartment are different (except for compartment 1). Using a model different than the sequential or parallel model is called a target analysis, and can only be done if the data allows for it and/or additional information is available from other techniques. The experiments described in this thesis are all analysed with all the above criteria in mind.

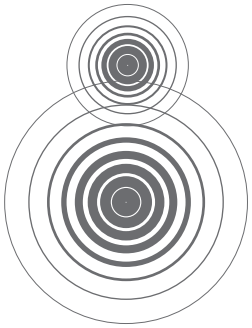
This thesis treats the above–described proteins and gives a unique insight in the dynamics occurring on an ultrafast time scale. Time–resolved infrared experiments are discussed in the first part of the thesis (chapter 2–5), where the results obtained on GFP, PYP, its E46Q mutant, and finally FixL are extensively discussed. Subsequently, polarized infrared experiments on PYP are presented in chapter 6, and a normal mode analysis on a model chromophore of PYP is presented in chapter 7. Finally, a complete summary of the work described in this thesis is given in chapter 8.

References

1. Camazine, S., Deneubourg, J.–L., Franks, N. R., Sneyd, J., Theraulaz, G. & Bonabeau, E. (2003) *Self–Organization in Biological Systems* (Princeton University Press, Princeton, New Jersey, USA).
2. Ohanian, H. C. (1988) *Physics* (W.W. Norton & Company, Inc., New York London.
3. Groenhof, G., Lensink, M. F., Berendsen, H. J. C., Snijders, J. G. & Mark, A. E. (2002) *PROTEINS: Structure, Function, and Genetics* 48, 202–211.
4. Groenhof, G., Lensink, M. F., Berendsen, H. J. C. & Mark, A. E. (2002) *PROTEINS: Structure, Function, and Genetics* 48.
5. Berg, J. M., Tymoczko, J. L. & Stryer, L. (2005) *Biochemistry* (W. H. Freeman, New York, USA).
6. Zimmer, M. (2002) *Chemical Reviews* 102, 759–781.
7. Tsien, R. Y. (1998) *Annual Review of Biochemistry 67, 509–* 544.
8. Agmon, N. (2005) *Biophysical Journal* 88, 2452–2461.
9. Yang, F., Moss, L. G. & Phillips, G. N. (1996) *Nature Biotechnology* 14, 1246–1251.
10. Shimomura, O. (1979) *Febs Letters* 104, 220–222.
11. Cody, C. W., Prasher, D. C., Westler, W. M., Prendergast, F. G. & Ward, W. W. (1993) *Biochemistry* 32, 1212–1218.
12. Kennis, J. T. M., Larsen, D. S., van Stokkum, I. H. M., Vengris, M., van Thor, J. J. & van Grondelle, R. (2004) *Proceedings of the National Academy of Sciences of the United States of America* 101, 17988–17993.
13. van Thor, J. J., Gensch, T., Hellingwerf, K. J. & Johnson, L. N. (2002) *Nature Structural Biology* 9, 37–41.
14. van Thor, J. J., Pierik, A. J., Nugteren–Roodzant, I., Xie, A. H. & Hellingwerf, K. J. (1998) *Biochemistry* 37**,** 16915–16921.
15. Striker, G., Subramaniam, V., Seidel, C. A. M. & Volkmer, A. (1999) *Journal of Physical Chemistry B* 103, 8612–8617.
16. Niwa, H., Inouye, S., Hirano, T., Matsuno, T., Kojima, S., Kubota, M., Ohashi, M. & Tsuji, F. I. (1996) *Proceedings of the National Academy of Sciences of the United States of America* 93, 13617–13622.
17. Chen, M. C., Lambert, C. R., Urgitis, J. D. & Zimmer, M. (2001) *Chemical Physics* 270**,** 157–164.
18. Litvinenko, K. L., Webber, N. M. & Meech, S. R. (2001) *Chemical Physics Letters* 346, 47–53.
19. Chattoraj, M., King, B. A., Bublitz, G. U. & Boxer, S. G. (1996) *Proceedings of the National Academy of Sciences of the United States of America* 93, 8362–8367.
20. Palm, G. J., Zdanov, A., Gaitanaris, G. A., Stauber, R., Pavlakis, G. N. & Wlodawer, A. (1997) *Nature Structural*

Biology 4, 361–365.
21. Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormo, M. & Remington, S. J. (1997) *Proceedings of the National Academy of Sciences of the United States of Americ*a 94, 2306–2311.
22. Lill, M. A. & Helms, V. (2002) *Proceedings of the National Academy of Sciences of the United States of America* 99, 2778–2781.
23. Vendrell, O., Gelabert, R., Moreno, M. & Lluch, J. M. (2006) *Journal of the American Chemical Society* 128**,** 3564–3574.
24. Wang, S. & Smith, S. C. (2007) *Physical Chemistry Chemical Physics* 9**,** 452–458.
25. Wang, S. & Smith, S. C. (2006) *Journal of Physical Chemistry B* 110, 5084–5093.
26. Heim, R., Prasher, D. C. & Tsien, R. Y. (1994) *Proceedings of the National Academy of Sciences of the United States of America* 91, 12501–12504.
27. Heim, R. & Tsien, R. Y. (1996) *Curr Biol* 6, 178–82.
28. Sansom, M. S. P., Adcock, C. & Smith, G. R. (1998) *Journal of Structural Biology* 121, 246–262.
29. Garczarek, F. & Gerwert, K. (2006) *Nature* 439, 109–12.
30. Faxen, K., Gilderson, G., Adelroth, P. & Brzezinski, P. (2005) *Natur*e 437**,** 286–289.
31. leCoutre, J. & Gerwert, K. (1996) *Febs Letters* 398, 333–336.
32. Frank, R. A. W., Titman, C. M., Pratap, J. V., Luisi, B. F. & Perham, R. N. (2004) *Science* 306, 872–876.
33. DeCoursey, T. E., Morgan, D. & Cherny, V. V. (2003) *Nature* 422, 531–534.
34. Hellingwerf, K. J., Hendriks, J. & Gensch, T. (2003) *Journal of Physical Chemistry A* 107, 1082–1094.
35. Sprenger, W. W., Hoff, W. D., Armitage, J. P. & Hellingwerf, K. J. (1993) *J Bacteriol* 175, 3096–3104.
36. Cusanovich, M. A. & Meyer, T. E. (2003) *Biochemistry* 42, 4759–4770.
37. Larsen, D. S. & van Grondelle, R. (2005) *Chemphyschem* 6, 828–837.
38. Changenet–Barret, P., Espagne, A., Charier, S., Baudin, J. B., Jullien, L., Plaza, P., Hellingwerf, K. J. & Martin, M. M. (2004) *Photochemical & Photobiological Sciences* 3, 823–829.
39. Larsen, D. S., Vengris, M., van Stokkum, I. H. M., van der Horst, M. A., de Weerd, F. L., Hellingwerf, K. J. & van Grondelle, R. (2004) *Biophysical Journal* 86, 2538–2550.
40. Hoff, W. D., Dux, P., Hard, K., Devreese, B., Nugteren–Roodzant, I. M., Crielaard, W., Boelens, R., Kaptein, R., van Beeumen, J. & Hellingwerf, K. J. (1994) *Biochemistry* 33,

13959–62.
41. Chosrowjan, H., Mataga, N., Shibata, Y., Imamoto, Y. & Tokunaga, F. (1998) *Journal of Physical Chemistry B* 102, 7695–7698.
42. Premvardhan, L. L., van der Horst, M. A., Hellingwerf, K. J. & van Grondelle, R. (2003) *Biophysical Journal* 84, 3226–3239.
43. Chosrowjan, H., Mataga, N., Nakashima, N., Imamoto, Y. & Tokunaga, F. (1997) *Chemical Physics Letters* 270, 267–272.
44. Changenet, P., Zhang, H., van der Meer, M. J., Hellingwerf, K. J. & Glasbeek, M. (1998) *Chemical Physics Letters* 282, 276–282.
45. Groot, M.–L., van Wilderen, L. J. G. W., Larsen, D. S., van der Horst, M. A., van Stokkum, I. H. M., Hellingwerf, K. J. & van Grondelle, R. (2003) *Biochemistry* 42, 10054–10059.
46. Heyne, K., Mohammed, O. F., Usman, A., Dreyer, J., Nibbering, E. T. J. & Cusanovich, M. A. (2005) *Journal of the American Chemical Society* 127**,** 18100–18106.
47. Larsen, D. S., van Stokkum, I. H. M., Vengris, M., van der Horst, M. A., de Weerd, F. L., Hellingwerf, K. J. & van Grondelle, R. (2004) *Biophysical Journal* 87, 1858–1872.
48. van Brederode, M. E., Gensch, T., Hoff, W. D., Hellingwerf, K. J. & Braslavsky, S. E. (1995) *Biophys J* 68, 1101–9.
49. de Philip, P., Soupene, E., Batut, J. & Boistard, P. (1992) *Mol Gen Genet* 235, 49–54.
50. Delgado–Nixon, V. M., Gonzalez, G. & Gilles–Gonzalez, M. A. (2000) *Biochemistry* 39, 2685–91.
51. Dioumaev, A. K. (1997) *Biophysical Chemistry* 67, 1–25.
52. Gilles–Gonzalez, M. A. & Gonzalez, G. (2005) *J Inorg Biochem 99, 1–*22.
53. Key, J. (2004) in *Ph.D Thesis, Department of Biochemisry and Molecular Biology* (The University of Chicago, Chicago, IL), Vol. Ph.D., pp. 186.
54. Schotte, F., Soman, J., Olson, J. S., Wulff, M. & Anfinrud, P. A. (2004) *Journal of Structural Biology* 147, 235–246.
55. Albrecht, A. C. (1961) *Journal of Molecular Spectroscopy* 6, 84–108.
56. Cramer, C. J. (2002) *Essentials of computational chemistry* (John Wiley & Sons Ltd., Chichester, England).
57. Scott, A. P. & Radom, L. (1996) *Journal of Physical Chemistry* 100, 16502–16513.
58. van Stokkum, I. H. M., Larsen, D. S. & van Grondelle, R. (2004) *Biochimica et Biophysica Acta–Bioenergetics* 1657, 82–104.
59. Jung, G., Wiehler, J. & Zumbusch, A. (2005) *Biophysical Journal* 88, 1932–1947.



The Pathway for Proton Transfer in Green Fluorescent Protein

02

L.J.G.W. van Wilderen, I.H.M. Van Stokkum, J.C. Arents, J.T.M. Kennis, K.J. Hellingwerf, R. van Grondelle, M.L. Groot (Submitted)

Abstract

The color of the fluorescent light emitted by Green Fluorescent Protein (GFP) depends on a light-induced proton transfer reaction occurring in a 'proton-wire' that is formed by the chromophore, a water molecule (W22), Ser205 and Glu222. Freshly synthesized GFP in the ground state contains a chromophore that is protonated (i.e. neutral) and absorbs in the near UV, whereas upon formation of its excited state (ES) the emitted fluorescence shifts to the green part of the spectrum.

We have studied the ES dynamics of GFP using ultrafast visible/mid-infrared pump-probe spectroscopy to elucidate the dynamics of proton transfer in the proton-wire. Model calculations, based on available crystal structures, have proposed pathways starting from the chromophore or starting from the end of the wire, i.e. from Glu222. Our data shows that, following excitation, rapid (3 ps) protonation of Glu222 occurs, prior to deprotonation of the chromophore. The fully anionic form appears bi-exponentially, with time constants of 8 and 165 ps, respectively. We conclude that the chain of proton transfer reactions starts at the acceptor end *of the wire. Our conclusions corroborate recent electronic structure calculations done by Vendrell O, Gelabert R, Moreno M, Lluch JM (2006) J Am Chem Soc 128: 3564-3574 and Wang S, Smith SC (2007) Phys Chem Chem Phys 9: 452-458, who found that pulling a proton over the wire, initiated by transfer of a proton from the serine residue to the glutamate, is energetically most favorable.*

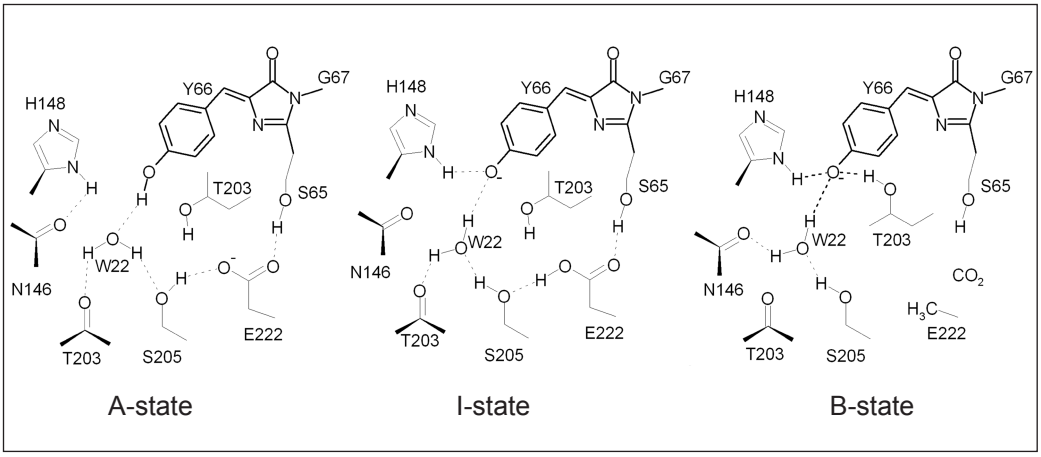
Introduction

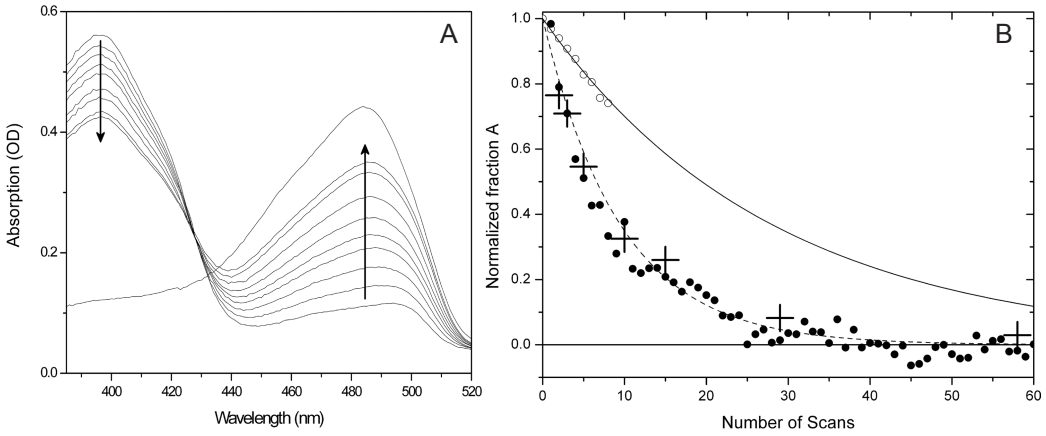
GFP from the jellyfish *Aequorea victoria* has become a popular fluorescent marker for molecular imaging in biology and medicine, due to its stability and excellent fluorescence properties (see (1-3) for a review). The chromophore of wild type (WT) GFP (and its more water-soluble derivative GFPuv) is auto-catalytically formed from three amino acid residues (Ser65-Tyr66-Gly67 (4)), absorbs in the near-UV, but fluoresces with a high quantum yield (0.8) in the green part of the spectrum. It is generally assumed that this large spectral shift is due to deprotonation of the chromophore in the excited state (5-7), since a kinetic isotope effect (KIE) is observed for the ultrafast rate of the spectral shift (5, 8). In addition, the identification in GFP's structure of a proton 'wire' formed by the hydrogen bonded chromophore, a water molecule (W22), a serine residue (S205) and a glutamate residue (E222; Figure 1), led to the proposal that a proton is relayed from the chromophore via the wire to the glutamate residue (9). This pathway was recently validated by the observation of the protonation of the glutamate side chain in time-resolved transient infrared experiments (10-12).

The absorption spectrum of GFP (Figure 2A) shows two bands, a main band near 400 nm, the A-state, in which the chromophore is protonated, and a minor band near 480 nm, referred to as the meta-stable B-state (see also Materials and Methods), in which the chromophore is in a deprotonated anionic form, with the nearby Glu222 decarboxylated (2, 13). The protein studied here is the GFPuv mutant, which contains some surface mutations but behaves similarly to WT-GFP. Upon excitation of the A-band, green fluorescence emitted from a state denoted I* appears on a picosecond time scale (5, 14)[1]. The fluorescent I*

Figure 1 **Schematic drawing of the chromophore-binding pocket of the protein in the A-, I-, and B-state (adapted from ref. Jung et al. (48), based on Brejc et al. (9)). The chromophore (bold lines) is formed from Ser65-Tyr66-Gly67, and is embedded in an extensive hydrogen bonding network (dashed lines). The A-state has a protonated Tyr66 and a negatively charged Glu222. The I-state has a deprotonated Tyr66 and a protonated Glu222. The B-state is formed by splitting off a CO_2 molecule upon excitation, leaving a decarboxylated Glu222 and a deprotonated Tyr66 (13).**

[1] *The I-state is also visible in the steady-state absorption spectrum as a band centered at 496 nm. This band is concentration dependent and probably caused by dimerization (15).*





state, supposedly the chromophore in deprotonated form, has been reported to be formed bi-phasically with time constants of ~2 ps and ~10 ps (5, 8, 14-16). These multi-phasic kinetics were assumed to reflect heterogeneity in the proton transfer rate, due to structural heterogeneities (5, 8, 14, 15), or to reflect the effect of vibrational cooling and the dissipation of energy into non-reactive vibrational modes (16), rather than to be associated with specific intermediates in the proton transfer pathway. In three time-resolved spectroscopic studies, performed in the mid-infrared part of the spectrum, protonation of Glu222 was observed to occur with the same kinetics. Therefore, deprotonation of the chromophore was concluded to be the rate limiting step (10-12).

The emission from the A* state is unusually broad, and in fact seems to reflect the presence of I* state emission already immediately upon excitation (5, 8, 16). This was previously discussed in terms of a non-vertical transition via overlapping wavefunctions of the A*- and I* -states (5, 15, 16). The observation of very early I* emission is hard to reconcile with chromophore deprotonation being the rate limiting step in the proton shuttle mechanism, since that implies that the subsequent reactions, involving transfer of a proton from the water molecule to the serine and from the serine to the glutamate, are even faster. Though so far all data showed a multi-phasic formation of I* and no intermediates in the proton transfer pathway were reported, this is different for the proton transfer reaction occurring upon the return of I* to the ground state. Back transfer of the proton to the chromophore is induced by the decay of the fluorescent anionic excited state to the anionic ground state, which occurs after about 3 ns (14, 17). Using visible pump-dump-probe spectroscopy an intermediate denoted I_2 was observed to be populated from the I-ground state in 3 ps and subsequently depopulated in 400 ps to the A-ground state (8), *i.e*. these proton transfer rates are much slower than the rates implied for the excited state proton transfer. We note that the

Figure 2A **Steady-state absorption spectra of GFPuv as a function of the number of pump-probe scans. The first spectrum is the fresh sample, the last one the fully converted one. Before and after every single scan the spectrum is measured. In panel B, the normalized fraction A-state, with reference to its amount in the fresh sample, is plotted as a function of the number of scans for: (i) the visible absorption peak at 398 nm (open circles) and (ii) the IR absorption peak at 1497 cm^{-1} (closed circles). For the latter, the signal between 1.1 and 1.2 ns was integrated. The plus signs (3 scans wide) represent the fraction A present in the sample, resulting from a simultaneous fit of 7 times the average of 3 scans (containing all wavenumbers) centered at 2, 3, 5, 10, 15, 29 and 59 scans. The solid and dashed lines are the concentration curves modeled with 3.5% and 10% permanently photoconverted sample per scan, respectively.**

first molecular dynamics study of the proton transfer processes in GFP by (18) did indeed suggest that chromophore deprotonation is the rate limiting step and that the subsequent transfers, from the water to the serine and from the serine to the glutamic acid, occurred in a few tens of femtoseconds. These calculations started off, however, with the chromophore in the deprotonated state. From a more recent study on the energetics of the proton transfer intermediates, by Vendrell *et al* and Wang *et al* (19-21), starting off with the chromophore in the protonated excited state, it appears that it would be energetically more favourable for the reaction to commence with the transfer of a proton from the serine to the glutamate, followed by the water to serine and chromophore to water proton transfers, so starting from the *end* of the proton wire. This opens the possibility that the observation of the protonation of the glutamate in the mid-IR spectroscopic studies represented not the end, but rather the beginning of the proton transfer reactions. We note that the observation of the red shifted stimulated emission in the visible part of the spectrum that represents the I^* state may not directly be taken as evidence of deprotonation of the chromophore. Electronic transitions of chromophores are very sensitive to several factors, such as charges in the vicinity of the chromophore (like the initially negatively charged glutamic acid), and the position of the emission maximum may also be affected by emission from 'hot' vibrational states. The vibrational modes of a chromophore on the other hand, as measured in the mid-infrared, are generally much better indicators of the electronic or redox state of a chromophore. In this study we have monitored the vibrational bands that are specific for the neutral and anionic state of the chromophore, using transient mid-infrared spectroscopy, to unequivocally determine the deprotonation rate of the chromophore. Following simultaneously the dynamics of the protonation marker band of Glu222 and the deprotonation marker bands of the chromophore in the infrared allows us to determine which of the proposed proton transfer processes described above is the correct one, and how fast the (proton) transfers proceed.

Resolving the pathway and dynamics of proton transfer in GFP has implications beyond the understanding of GFP itself. The water-Serine-Glutamate proton-wire might even be a recurring motif in other proteins. As proton transfer in GFP is initiated by the absorption of light, GFP can therefore serve as an important model system for biologically relevant proton transfer reactions involving proton wires. Proton transfer can therefore be studied without the complications caused by diffusion type dynamics. Proton wires play an important role in proteins, with functions ranging from pumping protons over membranes (22-25), to the catalytic function of enzymes (26, 27).

Measurement Method

Protein preparation

All experiments were performed with GFPuv. This derivative possesses several surface mutations (F99S, M153T and V163A) as compared with WT-GFP, but has WT UV/VIS absorption characteristics. Samples, dissolved in 10 mM Tris buffer (pH 8), were prepared as previously described (31) except that a pH gradient (pH 8-3.5) is used to elute the protein from the Ni column (HisTrap) instead of an imidazole gradient. For deuteration, samples were dissolved in D_2O-buffer, washed three times and left overnight in D_2O buffer to ensure complete H/D exchange. The sample consists of a highly concentrated protein solution pressed between two CaF_2 windows, separated by a Teflon spacer with a thickness of 6 μm, or with no spacer at all, depending on the probed region. The OD at 397 nm is about 0.9 at the start of the experiment. The ratio between the A and B peak (at 485) nm is about 7 for a freshly prepared sample.

Experimental setup

Visible absorption spectra are measured on a UV-VIS spectrometer (Lambda 40, Perkin-Elmer, Wellesley, MA, USA). The experimental setup (see (43) for a more detailed description) for the infrared difference absorption measurements consists of an integrated Ti:sapphire oscillator/ regenerative amplifier, operating at 1 kHz, and producing 0.8 mJ pulses of 85 fs (Hurricane, SpectraPhysics, Mountain View, CA, USA). The output of this laser is used to pump a commercial optical parametric generator and amplifier with difference frequency generation (TOPAS, Light Conversion, Vilnius, Lithuania), which results in a tunable output (2.5-10 μm) with a spectral width of ~200 cm^{-1}. A home-built HgCdT camera system placed behind a spectrograph is read out every shot at a repetition rate of 1 kHz and a sampling resolution of ~6 cm^{-1}. Another part of the Hurricane output is frequency doubled in a BBO crystal to generate the pump-pulses at 402 nm (~6 nm FWHM) with a power density of 200 nJ for excitation. A moveable delay line (double pass) makes it possible to increase the time-of-arrival-difference of the pump-and probe beams to 5 ns. The pump beam polarization is set with a Berek-rotator (New Focus 5540, San Jose, CA, USA) to magic angle with respect to the probe beam. A phase-locked chopper at 500 Hz ensures that every other shot the sample is excited and an absorbance difference spectrum can be calculated. To ensure a fresh spot for each laser shot, the sample is moved with a homebuilt Lissajous scanner. The setup is contained in a dry-air-purged box to reduce distortions of the infrared beam by absorption of water vapor.

Homebuilt software was used to collect data over the spectral window from 1062-1757 cm^{-1} in 5 partially overlapping windows. Every window was recorded with a freshly prepared sample and

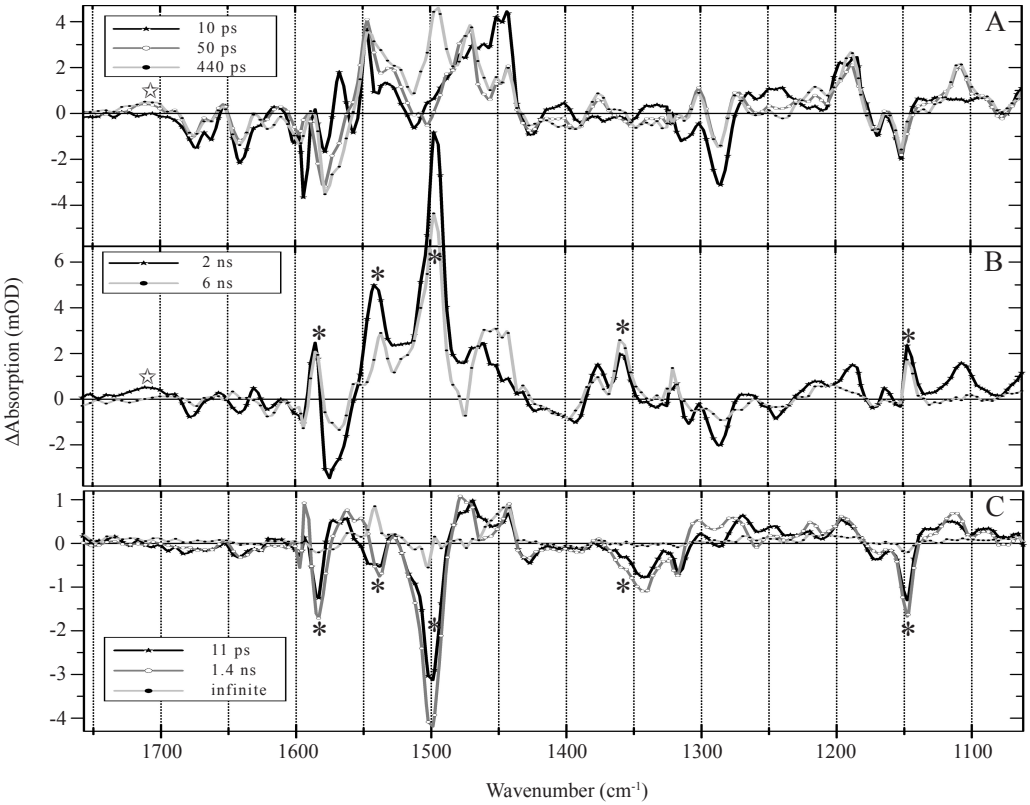


Figure 3 **EADS of GFP in D_2O, resulting from a global analysis of 7 sets of averages of 3 scans using a sequential model. Panel A and B represent kinetics of the A-state, panel C those of the B-state. The dynamics of the A-state are subdivided into panel A and B for clarity. The asterisks represent the markers for the deprotonated chromophore in the I-state, appearing with 50 ps in panel A (note the 1497 cm^{-1} band) and 440 ps in panel B, and disappearing instantaneously in the B-form (panel C). The open star marks the Glu222 band.**

at least measured twice. A complete data recording session of 60 scans took about 2 hours. A correction was made for the presence of a pre-time-zero offset due to thermal lens effects (44, 45). Because of the presence of a small perturbed free-induction decay before t = 0, no information faster than the instrument response (~200 fs) is extracted. The full H_2O- and D_2O datasets of the A and B-state had a noise level of 260 μOD.

Phototransformation of the sample

Upon photoexcitation of the A-state, GFP is photoconverted into the B-state. The B-form is formed with a low quantum yield and has a very long life time (days) (31, 46). Since we move our sample with a Lissajous scanner before each laser shot (43), each part of the sample is illuminated with only one or two laser pulses during one scan of the delay line. We have measured and characterized the gradual increase of the relative number of protein in the B-state, in order to obtain pure A-state dynamics. Two different data analysis methods are used. First, all data is divided into two parts to obtain an estimate for the A- and B-state kinetics; the first 3 scans represent the A-state, while the 58th until the 60th scan represent the decarboxylated B-state (resulting in Figure 3).

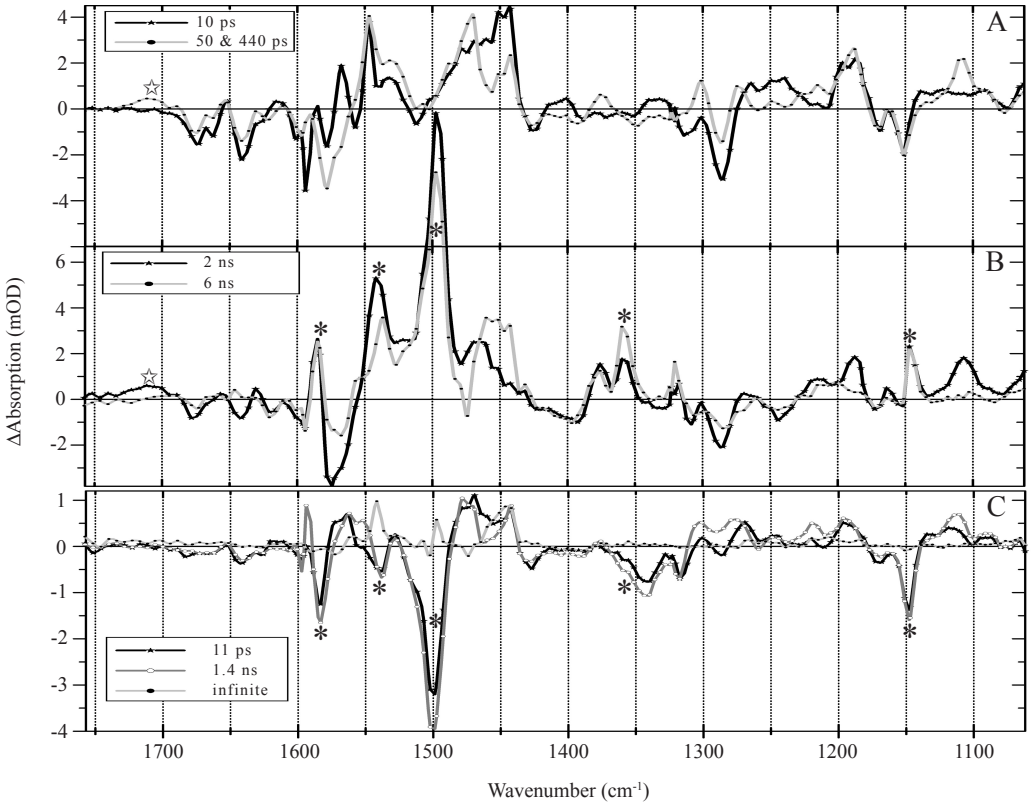


Figure 4 **SADS of GFP in D_2O, resulting from a target analysis of 7 sets of averages of 3 scans using the partially sequential/ parallel model (see Figure 5). Panel A and B represent the kinetics of the A-state, panel C those of the B-state. The dynamics of the A-state are subdivided into panel A and B for clarity. The asterisks represent the markers for the deprotonated chromophore in the I-state, appearing with 440 ps in panel B, and disappearing instantaneously in the B-form (panel C). The open star marks the Glu222 band.**

The second used method consists of performing a global analysis on 7 sets of averages of 3 scans, centered at 2, 3, 5, 10, 15, 29 and 59 scans, representing the partially converted sample (resulting in Figure 4 for D_2O and 8 for H_2O). The visible absorption spectrum of the sample is recorded after each scan (for the first 8 scans and the fully converted sample; Figure 2A). The stability of the B-state seems to depend on the excitation wavelength of GFP; 256 and 400–nm excitation (of the A state) forms an apparently irreversible B-state, while 476–nm excitation eventually leads to a full recovery (of the B-state) (31, 46). An initially present isosbestic point located at 428 nm is eventually lost, which has been suggested to be consistent with the photoconversion between two forms followed by destruction of the chromophore (38). It could be caused by a multi-photon process, in which an electron from the already excited (deprotonated) chromophore is ejected with a second photon, and causing the Glu222 to decarboxylate (13) and split off a CO_2 molecule with a low quantum yield. The altering ratio of the bands at 396 and 484 nm is then a measure for the transformation of the neutral chromophore into its anionic (decarboxylated) form. In Figure 2B we show, however, that the fractional change of absorption at 398 nm (represented by the open circles) is not a good indicator for the decarboxylation of the

protein. The induced absorption band at 1497 cm^{-1} in D_2O, typical for I* dynamics, is completely bleached (compare Figure 3B with 3C, and see Figure 7) and replaced by a negative signal at 1502 cm^{-1} after ~24 scans (not shown), whereas the absorption at 398 nm has only reduced by 58%. In Figure 2B (closed circles) the relative amplitude of the signal at 1497 cm^{-1} (integrated between 1.1-1.2 ns to increase the signal-noise ratio) is drawn as a function of time. Because of the presence of an underlying vibronic band of the B-state around 400 nm (5, 47), measurement of the apparent disappearance of the A-state at 398 nm leads to an overestimation of the fraction of A-state present. The true relative amount of A-fraction as monitored by the decrease of the absorption band at 1497 cm^{-1} can be modeled adequately by $A(n)=(1-c_{ex})^n$, where c_{ex} denotes the fraction of protein in the A-state which is converted into the B-state upon electronic excitation and n the number of scans. This formula describes the evolution of the A-state by subtracting the amount of photo-excited protein that is photo-transformed irreversebly. We find that c_{ex} is about 10% for the band at 1497 cm^{-1} (represented by the dashed line in Figure 2B), and 3.5% for the band at 398 nm (solid line in Figure 2B). Because the ratio A/B-state decreases from one scan to the next, we divide the scans into strategically chosen parts and analyze these data simultaneously with a model that is a linear combination of the kinetics of the A- and B-state. The average of 3 scans is taken to increase the signal-to-noise ratio. The simultaneous fitting method confirms the loss of ~10% of the neutral state per scan (plus-signs in Figure 2B). The average of the last three scans is virtually indistinguishable from the average of scans 15-60, indicating that the infrared-difference spectra are rapidly dominated by the anionic species.

Results

Neutral and anionic transient spectra in the 1062-1757 cm^{-1} region

Upon photoexcitation, GFP is converted from the neutral chromophore in the A-state (absorbing at 400 nm) to the anionic chromophore in the B-state (absorbing at 480 nm). The GFPuv samples in D_2O and H_2O are excited at 402 nm with an 85 fs laser pulse, and the induced absorption changes are probed in the mid–infrared fingerprint region from –16 ps to 5 ns. All collected data were subjected to a global analysis (in which all time traces have been fitted with the same set of kinetic parameters) using a sequential scheme with increasing lifetimes (Figure 6A of the supplementary information, from now on referred to as SI), to visualize the spectral evolution. An estimate of the time constants that describe the neutral and anionic states of the chromophore is obtained by fitting the average of the first three and last three

	Intact Glu222			Decarboxylated Glu222 (B-form)		
	H_2O	D_2O	KIE	H_2O	D_2O	KIE
A-form	3 (1) ps	10 (1) ps	~3	9 (2) ps	11 (2) ps	~1
	8 (2) ps	50 (6) ps	~6	1.2 (0.2) ns	1.4 (0.1) ns	~1
	165 (20) ps	440 (30) ps	~3	4.8 (0.8) ps	Long-lived	~1
I-form	880 (40) ps	2.1 (0.4) ns	~2			
	7.5 (2.5) ns	6.3 (1.0) ns	~1			

Table 1 **Global analysis parameters as a result of fitting the first three and last three scans of all data, representing the neutral and decarboxylated GFPuv, respectively, for different buffer conditions. Errors are shown in brackets. KIE is defined as the D_2O/H_2O ratio.**

scans, respectively (see SI). Five time constants (Table 1) were needed to obtain a satisfactory fit of the data for the A–state in D_2O (H_2O values in parentheses): 10 (3) ps, 50 (8) ps, 440 (165) ps, 2.1 (0.9) ns and 6.3 (7.5) ns. To increase the signal–to–noise, more scans of the partially converted sample are analyzed simultaneously. The time constants found with the analysis of the first three and last three scans are used as input parameters for the analysis of the sets of averages of the partially converted sample. For a more elaborate description of this data analysis method, the reader is referred to the Measurement methods section. Because the resulting absorption difference spectra for the first/last scan averages are virtually identical to those of the partially converted averages sample, only the latter results are shown. The evolution-associated difference spectra (EADS) for each of the lifetimes are shown in the upper two panels of Figure 3. Because the spectrum formed after 50 ps can be recognized to be a linear combination of the preceding and succeeding spectra, a target model (28) is applied to the data that allows for a biexponential formation of the 440 ps spectrum from the 10 ps species (Figure 6B of the SI). Essentially it represents a simpler model (containing 4 species with different spectral characteristics instead of 5), and has a fit quality that is virtually indistinguishable from the sequential model. The A* state thus forms the I*–state via a sequential step that is followed by a biexponential decay (with time constants of 50 ps and 440 ps with relative contributions of ~20 and ~80%, respectively). The species-associated difference spectra (SADS) for each of the A-state lifetimes are shown in Figure 3A and B. Analysis of the anionic form revealed that two lifetimes of 11 (9) ps, 1.4 (1.2) ns and a component that did not decay (4.8 ns) on the time scale of our experiment were sufficient to fit the data. The resulting EADS are shown in Figure 3C. The robustness of the fitting procedure is also demonstrated by near-identical difference spectra of the B-state for both fitting models (compare Figure 3C with 4C). Note that in the EADS/SADS negative bands are due to bleached ground state absorption and positive bands arise from product bands in the excited- or product states. The time constants describing the dynamics of the A-state are very similar to those observed by visible (pump-probe and fluorescence) spectroscopy

(5, 8, 14, 15, 29), and to those reported in earlier mid-IR studies on GFP in D_2O: 3-15, 49-75 ps and 1.7 ns (10, 12). Our results for the A-state of GFP are therefore consistent with these reports, except for the fact that we observe an additional component of 440 ps (165 ps in H_2O). This additional component most probably corresponds to the 120-200 ps component previously observed with fluorescence decay experiments in H_2O (7, 14, 15). The extra component was reported to be only present in the A* emission and not in the I* emission, and to display a moderate KIE (14). The values we obtain for the anionic form are similar to those reported by Lossau *et al*. (14) for the fluorescence decay of B* (values measured in H_2O in brackets): 12 (2.8) ps, and 1.2 ns (absent), although their second component (48 (20) ps) is not present in our data and we observe virtually no deuteration effect.

Isotope labeling and DFT-calculations (30, 31) on a model compound for the GFP chromophore have identified the following markers for the anionic chromophore: 1147 cm^{-1} (phenolic moiety of Tyr66 possibly mixed with the C_8-C_9 stretch frequency (31), 1341-1360 cm^{-1} and 1497/1582 cm^{-1} (phenolic moiety of Tyr66 (30)), and 1537 cm^{-1} (imidazolinone ring). All these bands, indicated with an asterisk in Figure 3B, appear dominantly (~80%) with the 440 ps time constant in the D_2O samples. Therefore, the two spectra of Figure 3B (and 4B), having 2- and 6-ns lifetimes, can be identified as the deprotonated excited state of the phenolate I^*-form and the deprotonated phenolate ground state I_2, respectively. The same markers are present as negative bands in the spectra in Figure 3C and 4C; indicating that the species present after 57 (and even already after 15) scans is indeed predominantly the anionic form of the chromophore.

The visible pump-probe-, fluorescence decay-, and visible pump-midIR probe data published so far, have all been analyzed with a heterogeneous model for the proton transfer, in which the neutral A* state decays bi-phasically into the I* state. As can be seen from inspection of the EADS spectra in Figure 3, our data are not consistent with such a model. The typical anion bands at 1583, 1537, 1497, 1360 and 1147 cm^{-1} appear (dominantly) with the 440 ps time constant. Furthermore, the dynamic changes associated with the 10 ps time constant are not similar to those of the 50 ps time constant: they show no hint of a common product state that is populated with the two initial time constants. For example, the 10 ps time constant is related to changes in the 1710 cm^{-1} region, the disappearance of a negative band at 1557 cm^{-1} and the appearance of a positive band at 1110 cm^{-1} (see Figure 7A of the SI). The main feature of the 50-ps process is the appearance of a positive band at 1493-1497 cm^{-1} (note that this is one of the characteristic markers for the anionic chromophore; it more than doubles in size after 440 ps). It seems therefore clear that the 10-ps and 50-ps time constants represent two physically distinct

processes, in spite of the fact that both components lead to a red shift of the emission in visible transient experiments to 508 nm (5, 7, 8, 16).

Discussion

Our results show that, in contrast to the commonly accepted view, several events take place before the anionic I* state appears. The KIE's of these (sequential) processes, ~3 for the first, ~6 for the second and ~3 for the third, indicate that each of these events is associated with proton transfer. To obtain further insight into their origins, we will discuss the corresponding midIR difference spectra in detail below and, with the aid of previous normal mode calculations (32, 33), isotopic labeling (30), Raman- (34, 35) and infrared spectroscopy (10, 12, 31, 36) of GFP protein and model chromophores, identify the underlying physical processes.

Due to the fact that earlier reports of time-resolved midIR difference spectra either did not fit all wavelengths (10, 11) or used a parallel model (12), we cannot directly compare our obtained (EADS/SADS) spectra. In addition, there is a discrepancy in the reported time constant for individual bands: In ref (12) the main anionic marker band at 1493 cm^{-1} (we find 1497 cm^{-1}) is reported to appear in 3 ns, *i.e.* it was observed only in the state identified as the anionic ground state of the chromophore and not in the excited state of the anionic chromophore, whereas we find it to appear with the 50 ps and 440 ps time constants. The I^* and I_2 difference spectra in Fig. 3B (and 4B) are overall very similar, while the corresponding spectra reported by Van Thor *et al.* are very different (12). More importantly, the I^*-A spectrum reported in (12) lacks all anionic markers. The EADS of the anionic chromophore (Figure 3C) show an instantaneous bleach at 1502 cm^{-1}, while the 1497 cm^{-1} positive band appears mainly with ~400 ps in the 'clean' A-state dynamics (for D_2O, see Figure 3B and 4B, for H_2O, see Figure 8 of the SI). Consequently, a gradual appearance of the B-state during the scan may have 'delayed' the appearance of the positive signal.

Glu222

The transients in the 1695-1720 cm^{-1} region can in general be assigned to carboxylic acid groups. The 1712 cm^{-1} band, which appears after 10 ps (marked by the open star in Figure 3A), also observed by Van Thor *et al.* and Meech *et al*. in their IR difference spectra (10-12), was assigned to the COOD mode of Glu222 in D_2O (10), based upon the fact that these signals were considerably smaller, delayed and slightly shifted in frequency in a mutant that had the glutamic acid replaced by an aspartic acid. The downshift of the signal to 1705 cm^{-1} in the I_2 state, occurring after 2 ns, has been interpreted as a weakening of (a) hydrogen bond(s) to

Glu222 (12). The appearance of the anionic B state absorption at 480 nm has been shown to be accompanied by decarboxylation of Glu222 (13). In our spectra we observe indeed that the signals at 1711 and 1705 cm^{-1} have disappeared when predominantly the anionic form of the protein is excited (note the flat SADS in the 1700 cm^{-1} region in Figure 3C and 4C), confirming the assignment of these bands to Glu222. The negative band at 1557 cm^{-1} which disappears in ~10 ps (black to dark grey SADS in Figure 3A and 4A) after excitation of the neutral chromophore has been assigned to the disappearance of the antisymmetric stretch of Glu222–COO^-, on the basis of its absence in the spectra of a model chromophore in solution, isotope labelling in the protein (11), and on the fact that it tracked the appearance of absorption of the COOD stretch of Glu222 (12). Here we observe that indeed the delayed bleaching of this band is absent when the anionic form of the chromophore is excited (Figure 3C, 4C and 8 of the SI). Therefore, we provide additional evidence for the protonation of the ionized glutamate as one of the key reactions in the GFP photocycle, and conclude that protonation of Glu222 occurs in ~10 ps in D_2O and in 3 ps in H_2O.

Chromophore modes and Gln69

The instantaneous bleach of a minor band at 1695 cm^{-1} (black SADS in Figure 3A) was earlier assigned to Gln69 as a response to the changed electron distribution in A* (12). The other bands in the spectra are assigned to chromophore modes. For example, the instantaneous bleach at 1673 cm^{-1} which decays mainly with the 2-ns time constant, after which a small 1673/1667 cm^{-1} band shift signal remains during the I_2 state, has been assigned to the bleaching of the C=O stretch mode of the chromophore (33). Most likely, formation of the excited state of the chromophore reorganizes the bond order of both chromophore rings, converting the C=O double bond to a single bond. One would expect therefore a C–O band to appear around 1260 cm^{-1} (37), and indeed a small positive feature is observed around 1250 cm^{-1} (also in H_2O and in the respective B-states, Figure 3, 4 and 8). The absence of these features in the spectra with the 2-ns and 6-ns lifetimes indicates that these signals decay with the ~440 ps time constant, implying that the bond order is at least partially restored upon deprotonation of the chromophore in the I* state. The feature at 1625+/1640– cm^{-1} is most likely mainly due to the exocyclic C=C bond that connects the two rings (11, 30, 38); the feature at 1595+/1604- cm^{-1} to a phenol ring mode (11, 30); and the two bands around 1450 cm^{-1} to phenol C–H deformation at 1439 cm^{-1} (30). The latter two bands have also been tentatively assigned to the low-frequency symmetric stretch vibrations of Glu222 COO^- (1456 and 1441 cm^{-1} in A_1^* and A_2^*, respectively (12)), but these bands are also present in our decarboxylated (Glu222-less, *i.e.* B-state)

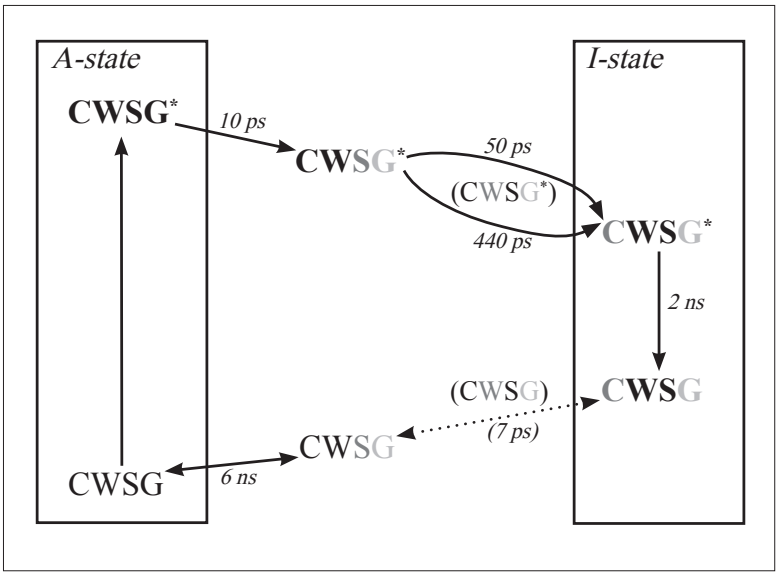


Figure 5 **Photophysical model with the time constants measured for the dynamics observed in our GFP-D_2O data, in combination with those reported in the visible pump-(dump-)probe data study by Kennis et al. (8). The ES of A is divided into three excited state proton transfer-steps. The four-letter code symbolizes the actors in the proton-wire (adapted from ref. (19)): C is the phenolic part of the chromophore (Tyr66), W is water molecule W22, S is Serine205 and G is Glu222. The bold states are the states identified in the current study in the IR. The 7-ps time constant between brackets is only observed in the visible pump-dump-probe data. The shades of grey denote the proton-state difference compared to the ground state: light grey means that it has received a proton, dark grey that is has lost one. Upon decay of the I* state, the proton transfer processes are reversed.**

spectra and therefore a phenol mode assignment seems more likely. As discussed earlier, the 1360 and 1150 cm^{-1} bands are phenol modes, sensitive to protonation, and the 1106 cm^{-1} mode may originate from C–H bending of the tyrosine ring (reported at 1105 cm^{-1} (39)) or possibly a C–O stretching vibration of a Threonine residue (39) or a combination of C–N stretching, C–H bending and N–H bending from the ring-structure of a Histidine (1095 cm^{-1} (39))

Decarboxylated B-state

The anionic markers of Figure 3C disappear instantaneously upon excitation, indicating that the B-state spectra are dominated by the anionic chromophore. Practically no KIE is observed for any of these components (Table 1). There is no indication from the spectral evolution for a reprotonation of the chromophore; rather the similarity of the EADS of the picoseconds and nanosecond components indicates a relaxation process. Vibrational cooling is most likely the cause for the observed spectral evolution. The EADS with infinite lifetime is similar to the 6-ns spectrum of the neutral form of the chromophore mixed with some leftover of B; it therefore is probably mainly due to a fraction of GFP with a neutral chromophore, following the normal A-photocycle. The B-form can only be formed if Glu222 is in the deprotonated form (13). Our infrared data show that Glu222 becomes protonated with the ~10-ps time constant, and the excited B-form can therefore only be formed from the initially excited state.

Proton pathway

The photophysical model that describes the data presented here is summarized in Figure 5. The four-letter code represents

the actors in the proton-wire (adapted from ref. (19)), with C, W, S and G symbolizing the phenolic part of the chromophore (Tyr66), water molecule W22, Serine205 and Glu222, respectively. The protonation state with respect to the ground state is also marked, with dark grey representing the donation of a proton, and light grey the acceptance of a proton. From our spectra it is clear that the Glu222-residue gets protonated in ~10 ps (D_2O; 3 ps in H_2O; it turns light grey in Figure 5), while the anionic markers of the chromophore are still absent. Therefore, it seems likely that the sequence of events as suggested by the energies resulting from the electronic structure calculations by Vendrell *et al*. and Wang *et all.* (19-21), in which first the protonation of the glutamic acid by the serine occurs, followed by the protonation of serine by the water and only then by the protonation of the water molecule by the chromophore, is correct. Note that the simulations of Wang *et al.* favor a more concerted mechanism while still distinct steps are found (20, 21). It is tempting to identify the intermediate that appears after 50 ps (light grey EADS spectrum in Figure 3A) as the state in which the water has donated a proton to the serine. The spectral changes associated with the 50 ps time constant are in fact fairly modest, the main spectral change is the partial appearance of the 1497 cm^{-1} mode, which is a marker for the deprotonation state of the phenol ring. Since all the other markers do not appear with this time constant, it can be speculated that the appearance of the 1497 cm^{-1} band signals a change in the hydrogen bond network of the phenolic ring, and this could indeed be consistent with the deprotonation of the water molecule. An alternative interpretation could be that the 440-ps EADS actually is a superposition of the 50-ps and 2-ns EADS. The alternative model with one spectrum less (presented in Figure 5 and 6B) results in similar overall fit-quality, and seems therefore a more likely candidate for the true course of events.

The appearance of the complete set of anionic band markers with the 50/440-ps process provides a strong indication that this is the time scale that the proton is transferred from the chromophore, most likely to the water molecule. The suggested sequence of events is in agreement with the observation of a KIE for each of the steps.

The interpretation of previously published visible transient (8) data, then changes however. Our interpretation of the femtosecond-infrared spectra implies that the red-shift of the emission spectrum of GFP occurs with the protonation of Glu222 (or the deprotonation of the serine residue), and not with the deprotonation of the chromophore. We propose therefore that it is the neutralization of the negative charge on Glu222 in the protein pocket (note that besides Glu222 also Arg96 carries a local (positive) charge) that causes the observed red-shift of the chromophore emission. The presence of charges within the chromophore-binding pocket

has been demonstrated to change the absorption spectrum of members of the rhodopsin-family. In general, interactions between the chromophore and the surrounding residues that stabilize the ES, will induce a red-shift, and interactions that stabilize the ground state a blue-shift (40). The human rhodopsin family absorbs in the visible part of the spectrum from the blue (~425 nm) to the red (~560 nm). This enormous 'opsin'-shift (ΔE= ~5700 cm^{-1}) is explained by direct interactions of dipolar residues with the ground state of the chromophore (shifting it from green to blue) and longer-range interactions combined with the change in electric dipole moment (10-15 D (41)) upon electronic excitation (shifting it to the red) (42). The red-shift observed for GFP (ΔE= ~5500 cm^{-1}) is similar to the 'opsin'-shift and could therefore very well be caused by the changed interactions between the dipolar environment and the excited-state charge distributions, as a result of the protonation of the glutamate residue. Thus, the chromophore emission maximum seems to be more sensitive to the changed local charges and changes in the hydrogen bond network rather than to the difference between the anionic and protonated state. In other words, we suggest that the color tuning in GFP is very simlar to that in the opsin-family, but, remarkably, is induced by the excited-state proton transfer reaction. To show that indeed visible pump-probe data of GFP may be equally well be analyzed with a sequential model as we used for the midIR data, we refitted the visible pump-probe data reported by Kennis et al (8), see Figure 9 of the SI. As expected, the emission when fitted with a sequential model now shows a more dual nature, with the main part of the emission shifting to the red in 3 ps (in H_2O) and a minor part in ~10 ps. Possibly, the protonation of the glutamic acid residue causes an equilibrium between the proton located on the donor water molecule and on the serine acceptor molecule (states denoted C**W**SG and CW**S**G in Figure 5) rather than a pure state in which the proton 'hole' is located on the serine residue. Since we earlier concluded that most likely the combined effect of the protonation of the Glu and the changed hydrogen bond interaction with the water molecule causes the red shift of the chromophore's emission, the creation of such an equilibrium state may cause an early partial shift of the chromophore's emission and result in the observed dual emission spectra (5, 8, 16). Note that the corresponding equilibrium constant cannot be estimated simply from the maxima of the emission wavelengths since it has to be taken into account that in the first tens of ps the emission is still quite hot, and that the emission is to chemically different ground states, therefore the exact energy level of the involved excited states cannot be estimated from the emission wavelengths. The subsequent deprotonation of the chromophore, in ~440 ps in D_2O and ~165 ps in H_2O data, does apparently not lead to a further change in electronic emission intensity or frequency.

Conclusions

We propose that the changed dipole moment in the protein pocket, due to the formation of the chromophore-excited state, causes protonation of Glu222, most likely by the serine. In turn, the protonation of Glu222 forms the driving force for a proton transfer reaction from the water to the serine and a final deprotonation of the chromophore. The latter two processes appear to occur in concerted fashion, since we can describe our data very well without explicitly assuming the population of the state where the water is deprotonated. The present experiments show that the role of the protein in the function and color tuning of GFP is more intricate than previously assumed, and is possibly more similar to that of the opsin-family. Ultimately this can be used to tweak the spectroscopic properties of this widely used fluorescent label and increase its utility in future applications in studies of molecular cell-biology.

Acknowledgements

The authors thank P. Bier for his help in the preparation of the sample, K. Mullen for her help with the data analysis, and Prof Holger Lill for critically reading the manuscript. This research was supported by The Netherlands Organization for Scientific Research through the Dutch Foundation for Earth and Life Sciences (Investment Grant 812.08.001 and Molecule to Cell-grant 805.47.123 to K.J.H. and fellowship 834.01.002 to M.-L.G.).

References

1. Zimmer, M. (2002) *Chemical Reviews* 102, 759-781.
2. Tsien, R. Y. (1998) *Annual Review of Biochemistry* 67, 509-544.
3. Remington, S. J. (2006) *Curr Opin Struct Biol* 16, 714-21.
4. Cody, C. W., Prasher, D. C., Westler, W. M., Prendergast, F. G. & Ward, W. W. (1993) *Biochemistry* 32, 1212-1218.
5. Chattoraj, M., King, B. A., Bublitz, G. U. & Boxer, S. G. (1996) *Proceedings of the National Academy of Sciences of the United States of America* 93, 8362-8367.
6. Heim, R., Prasher, D. C. & Tsien, R. Y. (1994) *Proceedings of the National Academy of Sciences of the United States of America* 91, 12501-12504.
7. Kummer, A. D., Wiehler, J., Rehaber, H., Kompa, C., Steipe, B. & Michel-Beyerle, M. E. (2000) *Journal of Physical Chemistry B* 104, 4791-4798.
8. Kennis, J. T. M., Larsen, D. S., van Stokkum, I. H. M., Vengris, M., van Thor, J. J. & van Grondelle, R. (2004) *Proceedings of the National Academy of Sciences of the United States of America* 101, 17988-17993.
9. Brejc, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormo, M. & Remington, S. J. (1997) *Proceedings of the National Academy of Sciences of the United States of America* 94, 2306-2311.
10. Stoner-Ma, D., Jaye, A. A., Matousek, P., Towrie, M., Meech, S. R. & Tonge, P. J. (2005) *Journal of the American Chemical Society* 127, 2864-2865.
11. Stoner-Ma, D., Melief, E. H., Nappa, J., Ronayne, K. L., Tonge, P. J. & Meech, S. R. (2006) *Journal of Physical Chemistry B* 110, 22009-22018.
12. van Thor, J. J., Zanetti, G., Ronayne, K. L. & Towrie, M. (2005) *Journal of Physical Chemistry B* 109, 16099-16108.
13. van Thor, J. J., Gensch, T., Hellingwerf, K. J. & Johnson, L. N. (2002) *Nature Structural Biology* 9, 37-41.
14. Lossau, H., Kummer, A., Heinecke, R., PollingerDammer, F., Kompa, C., Bieser, G., Jonsson, T., Silva, C. M., Yang, M. M., Youvan, D. C. & Michel-Beyerle, M. E. (1996) *Chemical Physics* 213, 1-16.
15. van Stokkum, I. H. M., Gobets, B., Gensch, T., van Mourik, F., Hellingwerf, K. J., van Grondelle, R. & Kennis, J. T. M. (2006) *Photochemistry and Photobiology* 82, 380-388.
16. Winkler, K., Lindner, J. R., Subramaniam, V., Jovin, T. M. & Vohringer, P. (2002) *Physical Chemistry Chemical Physics* 4, 1072-1081.
17. Perozzo, M. A., Ward, K. B., Thompson, R. B. & Ward, W. W. (1988) *Journal of Biological Chemistry* 263, 7713-7716.
18. Lill, M. A. & Helms, V. (2002) *Proceedings of the National Academy of Sciences of the United States of America* 99,

2778-2781.
19. Vendrell, O., Gelabert, R., Moreno, M. & Lluch, J. M. (2006) *Journal of the American Chemical Society* 128, 3564-3574.
20. Wang, S. & Smith, S. C. (2007) *Physical Chemistry Chemical Physics* 9, 452-458.
21. Wang, S. & Smith, S. C. (2006) *Journal of Physical Chemistry B* 110, 5084-5093.
22. Sansom, M. S. P., Adcock, C. & Smith, G. R. (1998) *Journal of Structural Biology* 121, 246-262.
23. Garczarek, F. & Gerwert, K. (2006) *Nature* 439, 109-12.
24. Faxen, K., Gilderson, G., Adelroth, P. & Brzezinski, P. (2005) *Nature* 437, 286-289.
25. leCoutre, J. & Gerwert, K. (1996) *Febs Letters* 398, 333-336.
26. Frank, R. A. W., Titman, C. M., Pratap, J. V., Luisi, B. F. & Perham, R. N. (2004) *Science* 306, 872-876.
27. DeCoursey, T. E., Morgan, D. & Cherny, V. V. (2003) *Nature* 422, 531-534.
28. van Stokkum, I. H. M., Larsen, D. S. & van Grondelle, R. (2004) *Biochimica et Biophysica Acta-Bioenergetics* 1657, 82-104.
29. Leiderman, P., Ben-Ziv, M., Genosar, L., Huppert, D., Solntsev, K. M. & Tolbert, L. M. (2004) *Journal of Physical Chemistry B* 108, 8043-8053.
30. He, X., Bell, A. F. & Tonge, P. J. (2002) *Journal of Physical Chemistry B* 106, 6056-6066.
31. van Thor, J. J., Pierik, A. J., Nugteren-Roodzant, I., Xie, A. H. & Hellingwerf, K. J. (1998) *Biochemistry* 37, 16915-16921.
32. Yoo, H. Y., Boatz, J. A., Helms, V., McCammon, J. A. & Langhoff, P. W. (2001) *Journal of Physical Chemistry B* 105, 2850-2857.
33. Esposito, A. P., Schellenberg, P., Parson, W. W. & Reid, P. J. (2001) *Journal of Molecular Structure* 569, 25-41.
34. Bell, A. F., He, X., Wachter, R. M. & Tonge, P. J. (2001) *Biochemistry* 40, 8619-8619.
35. Bell, A. F., He, X., Wachter, R. M. & Tonge, P. J. (2000) *Biochemistry* 39, 4423-4431.
36. van Thor, J. J., Georgiev, G. Y., Towrie, M. & Sage, J. T. (2005) *Journal of Biological Chemistry* 280, 33652-33659.
37. Barth, B. (2000) *Progress in Biophysics & Molecular Biology* 74, 141-173.
38. Bell, A. F., Stoner-Ma, D., Wachter, R. M. & Tonge, P. J. (2003) *Journal of the American Chemical Society* 125, 6919-6926.
39. Wolpert, M. & Hellwig, P. (2006) *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy* 64, 987-1001.
40. Kloppmann, E., Becker, T. & Ullmann, G. M. (2005) *Proteins-Structure Function and Bioinformatics* 61, 953-965.
41. Mathies, R. & Stryer, L. (1976) *Proceedings of the National Academy of Sciences of the United States of America* 73,

2169-2173.
42. Kochendoerfer, G. G., Lin, S. W., Sakmar, T. P. & Mathies, R. A. (1999) *Trends in Biochemical Sciences* 24, 300-305.
43. Groot, M.-L., van Wilderen, L. J. G. W., Larsen, D. S., van der Horst, M. A., van Stokkum, I. H. M., Hellingwerf, K. J. & van Grondelle, R. (2003) *Biochemistry* 42, 10054-10059.
44. Marcano, A. O., Rodriguez, L. & Alvarado, Y. (2003) *Journal of Optics A: Pure and Applied Optics* 5, S256-S261.
45. Kurian, A., Unnikrishnan, K. P., George, D. S., Gopinath, P., Nampoori, V. P. N. & Vallabhan, C. P. G. (2003) *Spectrochimica Acta Part A* 59, 487-491.
46. Striker, G., Subramaniam, V., Seidel, C. A. M. & Volkmer, A. (1999) *Journal of Physical Chemistry B* 103, 8612-8617.
47. Bonsma, S., Purchase, R., Jezowski, S., Gallus, J., Konz, F. & Volker, S. (2005) *Chemphyschem* 6, 838-849.
48. Jung, G., Wiehler, J. & Zumbusch, A. (2005) *Biophysical Journal* 88, 1932-1947.

Supplementary Information

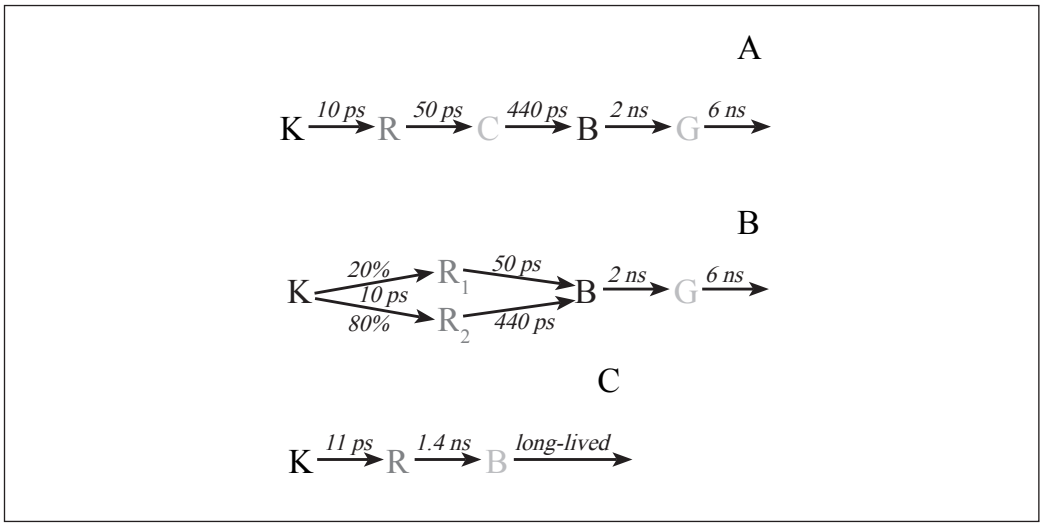


Figure 6 **Kinetic models used to describe the A-state (panels A and B) and the B-state dynamics (panel C). A fully sequential scheme is shown in panel A and a partially biexponential decay from R to B is shown in panel B. The two species R_1 and R_2 are spectrally identical, but form B with different yields and times. The states K, R, C, B, and G have a shade of grey that corresponds to the shade of grey of the species drawn in Figure 3A/B and 4A/B. The B-state is best described with a three-component sequential model (panel C). All time constants shown result from the analysis of GFPuv in D_2O.**

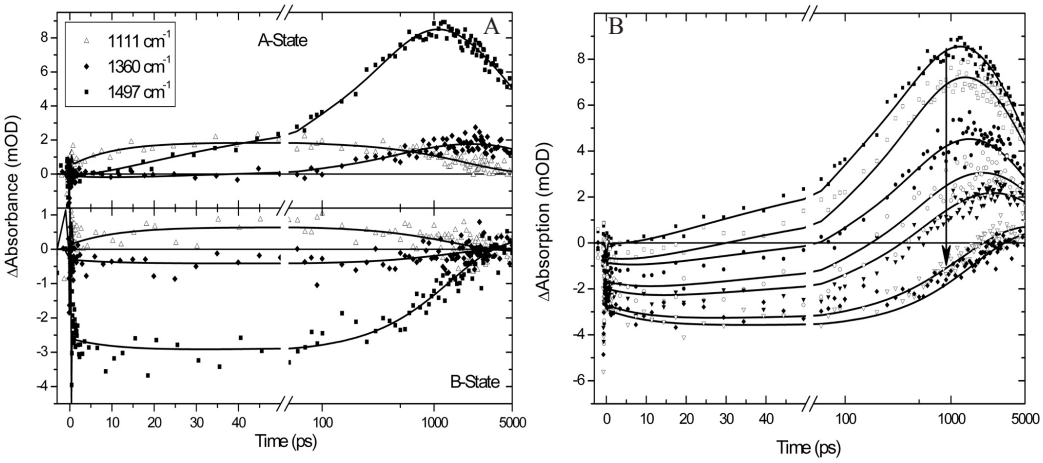


Figure 7 **Selection of time traces of GFP in D_2O. The left-top and left-bottom panels (A) show the time traces for the same wavelengths. The top panel represents the A-state and the bottom the (predominant) B-state, represented by the average of the first and last 3 scans. The different dynamics in the traces are clearly visible. In the right panel (B) 7 sets of averages of 3 scans of the raw data recorded at 1497 cm^{-1} are shown. The fits to the data that result from a target analysis (with the model drawn in Figure 5) are also shown. The average time traces centered at 2, 3, 5, 10, 15, 29 and 59 scans decrease with increasing number of scans (in symbols going from squares, open squares, spheres, open spheres, triangles, open triangles to diamonds, respectively). All panels have a time-axis that is linear up to 50 ps and logarithmic thereafter. Note that a partially converted sample will thus show an apparent delay of the formation of the 1497 cm^{-1} signal (i.e. when the formation of the decarboxylated state is not taken into account).**

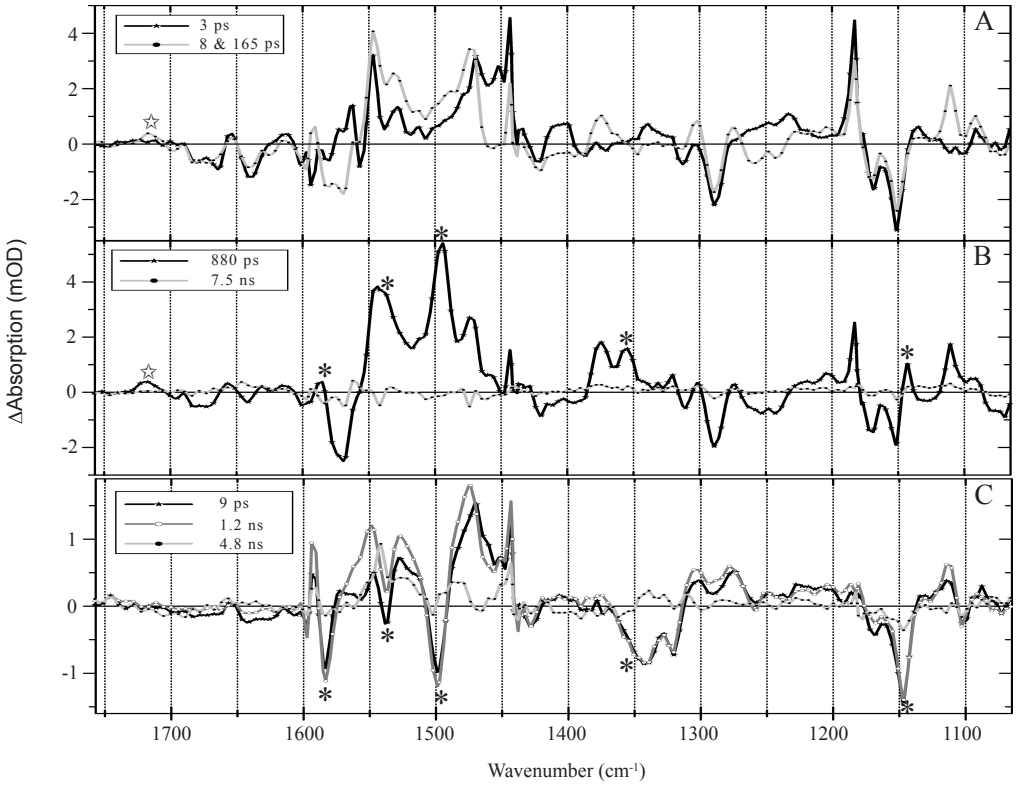


Figure 8 **Species associated difference spectra of GFP in H_2O resulting from a target analysis of 7 sets of averages of 3 scans (analogous to the D_2O analysis) with the partially sequential/parallel model (see Figure 5), representing the A-state (panel A and B) and B-state kinetics (panel C). The dynamics of the A-state are subdivided into panel A and B for clarity reasons. The asterisks represent the markers for the deprotonated chromophore in the I-state, appearing with 8 and 165 ps in panel B, and disappearing instantaneously in the B-form (panel C). The open star marks the Glu222 band.**

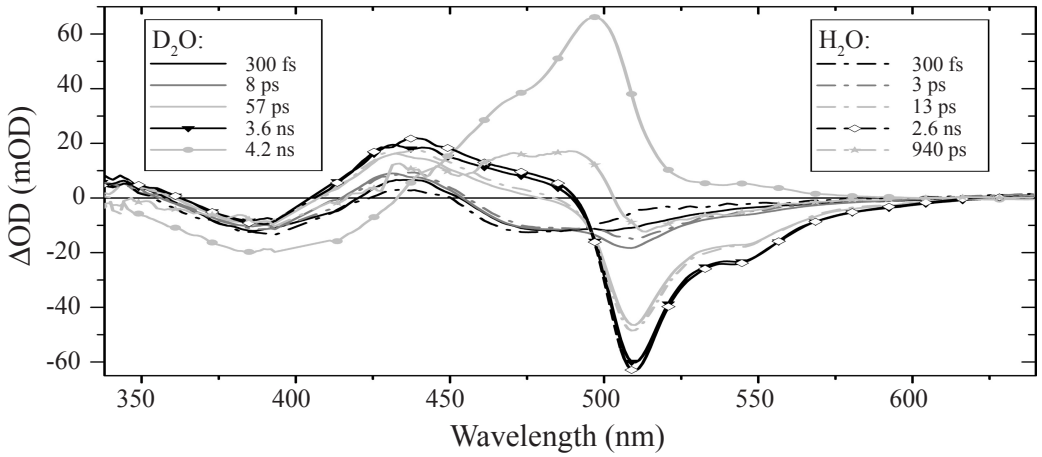
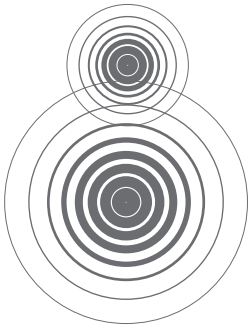


Figure 9 **Evolution associated difference spectra in the visible part of the spectrum of GFP in D_2O (solid lines) and H_2O (dot-dashed lines) resulting from a global analysis with a sequential model. The data presented here is based on the same data as the one in ref. (8). The last two components are reversed in H_2O to demonstrate the reversed kinetics. The light grey EADS (spheres and stars) represents the I_2 state CWSG, the black triangles and diamonds the I state CWSG.**



Initial Steps of Signal Generation in Photoactive Yellow Protein revealed with Femtosecond Mid-infrared Spectroscopy

03

Based on: M.L. Groot, L.J.G.W. van Wilderen, D.S. Larsen, M.A. van der Horst, I.H.M. Van Stokkum, K.J. Hellingwerf, R. van Grondelle Biochemistry 2003 42 (34) 10054-10059, 2003. 10.1021/bi034878p S0006-2960(03)04878-5.

Abstract

Photoactive Yellow Protein (PYP) is a bacterial blue-light sensor that induces Halorhodospira halophila to swim away from intense blue light. Light absorption by PYP's intrinsic chromophore, p-coumaric acid, leads to the initiation of a photocycle that comprises several distinct intermediates. Here we describe the initial structural changes of the chromophore and its nearby amino acids, using visible pump/mid-infrared probe spectroscopy. Upon photoexcitation, the trans bands of the chromophore are bleached and shifts of the phenol-ring bands occur. The latter are ascribed to charge translocation, which probably plays an essential role in driving the trans to cis isomerization process. We conclude that breaking of the hydrogen bond of the chromophore's C=O group with amino acid Cys69 and formation of a stable cis ground state occur in ~2 ps. Dynamic changes also include rearrangements of the hydrogen bonding network of the amino acids around the chromophore. Relaxation of the coumaryl tail of the chromophore occurs in 0.9-1 ns, which event we identify with the I_0 to I_1 transition observed in visible spectroscopy.

Introduction

Activation of PYP by blue light induces a negative phototactic response that enables the bacterium *Halorhodospira halophila* to swim away from harmful exposure to intense blue light (see for a review ref. *1*). Due to its rich and complex photocycle and its excellent (photo)chemical stability, PYP has become a model system for the study of biological signal generation in photoreceptor proteins. Light–induced signal generation in biology involves the amplification of an initially small configurational change generated in an active site, into a conformational change of a larger scale that converts the protein into its signalling state. The initial configurational change in the photocycle of PYP (*2*), i.e. *trans/cis* isomerization of its intrinsic chromophore (*p*–coumaric acid (*3,4*) depicted in Figure 1), is similar to that in other photosensors, like (bacterio)rhodopsin and (bacterio)phytochromes. A key characteristic of the PYP photocycle is a partial unfolding of the protein (*4*) on a sub–millisecond time scale, triggered by the absorption of a photon.

The PYP photocycle has been characterized by visible transient spectroscopy: The excited state of the chromophore decays multi–exponentially in a few picoseconds to a red shifted intermediate, denoted I_0 and absorbing around 500 nm (*6*–*11*), which is followed by the formation of an intermediate absorbing at 480 nm in 1–3 ns (*6*–*11*), denoted I_1 (also called pR or PYP_L). A blue shift of the chromophore absorption due to protonation (*13, 14*) occurs after ~0.3 ms, accompanied by a partial unfolding of the protein (pB or PYP_M state, *4,14*); this is probably the signalling state of PYP. Upon deprotonation and reisomerization of the chromophore, the ground state is recovered on a ~200 millisecond time scale. The structural changes of the chromophore and in the protein occurring during the photocycle have been studied with a wide range of techniques, like time–resolved X–ray crystallography (*15,16*) and time–resolved FTIR spectroscopy (*17*), from ~10 ns after excitation onwards, and by the application of these techniques on photocycle intermediates trapped by lowering the temperature (*18,19*). It was concluded that a few nanoseconds after light absorption, i.e. when PYP is in the I_1 state, important structural changes have already occurred. The chromophore has adopted the *cis* configuration (*15*–*18*), and changes in the hydrogen bonding network and structural rearrangements of nearby amino acids are observed. The chromophore phenolate ring has moved only slightly in the new configuration, and its hydrogen bonds with the nearby E46 and Y42 residues remain intact (*16*). Together with the covalent bond with Cys69, this limits the freedom of motion of the chromophore appreciably; therefore isomerization can take place only by simultaneous rotation around the dihedral angle of multiple bonds (*e.g.* $\theta_{7=8}$ and θ_{4-7}, as has been suggested previously, *20,21*). However, the time scale of these initial events

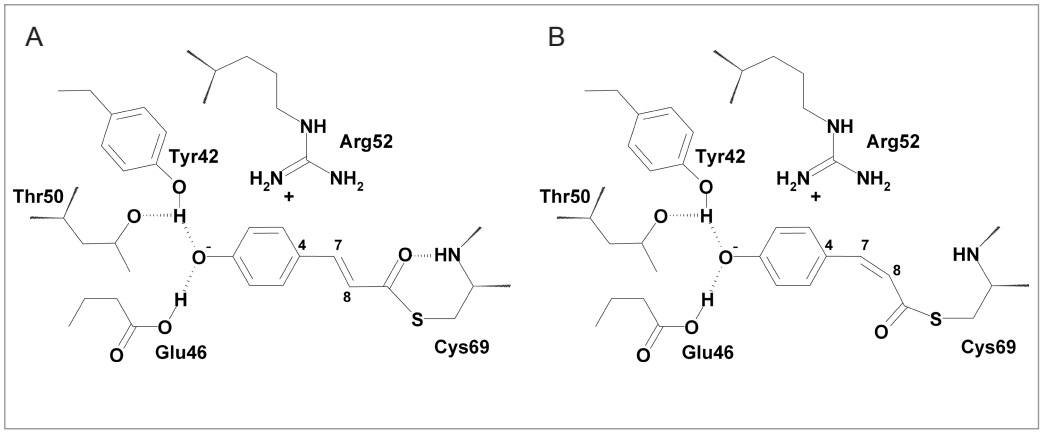


Figure 1 **The active site of PYP, depicting the chromophore covalently linked to Cys69, and the amino acids Glu46, Tyr42, Thr50 and Arg 52 near the phenolate ring of the chromophore, in the ground state (A) and in the I_1 state (B).**

has not yet been identified.

In the present study, we provide for the first time the 'missing link' between ultrafast optical information on the one hand and 'slower' structural information on the other by applying mid–infrared transient absorption spectroscopy with ~200 fs time resolution. In contrast to UV/Vis spectroscopy, vibrational spectroscopy is sensitive to precise structural configurations and can even have a sub–optical cycle time resolution of ~10 fs if a coherent–emission technique is used (*22*). Our data for the first time yields information on the structural changes that occur in the excited state, and in the first two ground state intermediates I_0 and I_1.

Materials and Methods

PYP was prepared as described previously (*23*). The sample was placed between two 2–mm thick CaF_2 plates separated by a 6 micron Teflon spacer. The use of concentrated PYP solutions (OD_{446} ~ 1.0) and thin spacers diminished the presence of the intense water absorption bands in the IR absorption spectrum.

The experimental setup consists of an integrated Ti:sapphire oscillator–regenerative amplifier laser system (Hurricane, SpectraPhysics) operating at 1 kHz, 800 nm, producing 85 fs pulses of 0.8 mJ. A portion of this 800–nm light was used to pump a noncollinear optical parametric amplifier to produce the excitation pulses with center wavelength of 475 nm, i.e. at the red edge of the absorption band, to achieve excitation into the $\upsilon = 0$ level for the 1000–2000 cm^{-1} modes. The excitation pulses had a duration (uncompressed) of about 60 fs. They were attenuated to 300 nJ to reduce local heating of the sample, and focused with a 20 cm lens into the sample. A second portion of the 800–nm light is used to pump an optical parametric generator and amplifier with difference frequency generator (TOPAS, Light Conversion) to produce the midIR probe pulses. The probe pulses were attenuated to an intensity of about 1 nJ and were spatially overlapped with the

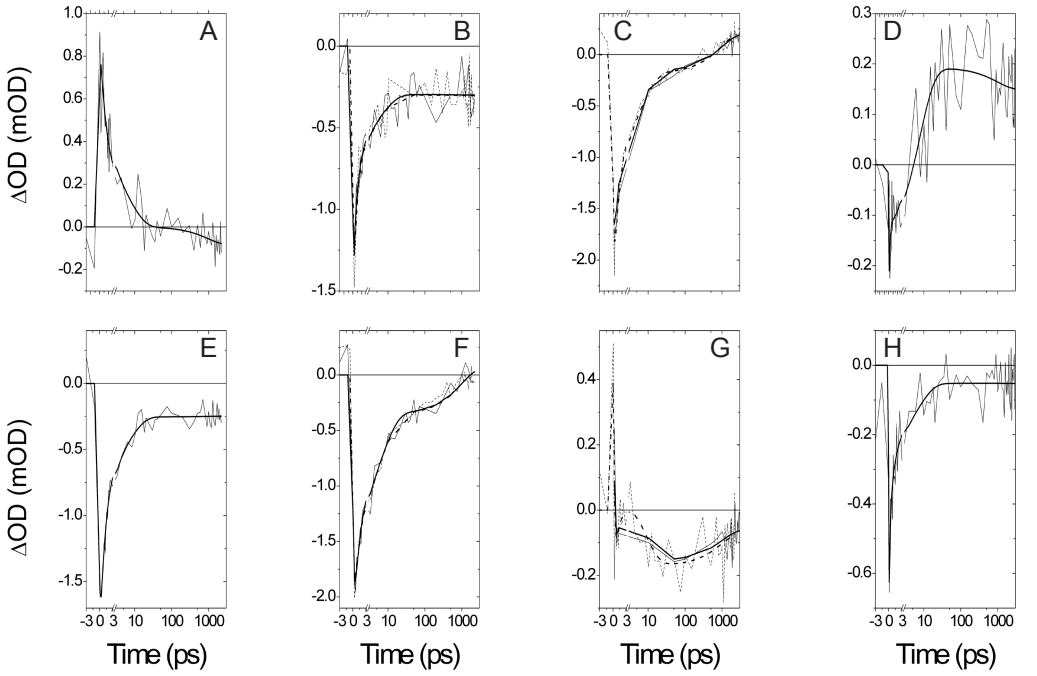


Figure 2 **Eight of the 120 time traces collected over the 1100–1850 cm^{-1} range upon excitation of the chromophore at 475 nm with a 60 fs pulse, plotted on a linear scale up to 3 ps and on a logarithmic scale for later delay times (up to 3 ns). The y–axis is in mOD units. Panels A-G represent the signals measured at 1142, 1268, 1487, 1662, 1163, 1310, 1515, and 1735 cm^{-1}, respectively. The smooth solid line in each panel represents the fit to the data; some figures show data from two different experiments, indicated by the different line styles. Before time zero a small perturbed free induction decay is observed.**

excitation beam in the sample. The cross correlation of the visible and IR pulses was measured in GaAs to be about 180 fs.

After overlap in the sample, the mid–IR probe pulses were dispersed in a spectrograph and imaged onto a 32–element MCT detector (Infrared Associates). The signals of the detector array were amplified (Infrared Associates) and fed into 32 home–built integrate–and–holds which were read out every shot with a National Instruments acquisition card (PCI6031E). The polarization of the excitation pulse was set to the magic angle (54.7^{o} with respect to the IR probe pulses. A phase–locked chopper operating at 500 Hz was used to insure that every other shot the sample was excited and the change in transmission and hence optical density could be measured. To ensure a fresh spot for each laser shot the sample was moved by a home–built Lissajous scanner. In a single experiment a spectral probe window of about 200 cm^{-1} was covered, so five partly overlapping regions were measured between 1850 and 1100 cm^{-1}. Experiments were repeated at least two times.

Results and Discussion

Figure 2 shows time traces collected over the 1100–1850 cm^{-1} region upon excitation of the PYP chromophore at 475 nm. A global analysis of the data revealed that three lifetimes are present in our data: 2 ps, 9 ps, 0.9–1 ns, and a non–decaying component (>10 ns). Before t = 0 a small perturbed free–induction decay signal

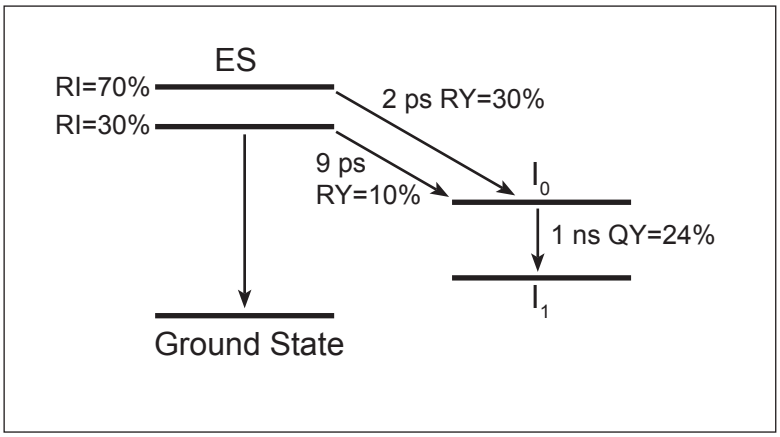


Figure 3 **Schematic depiction of the first part of the PYP photocycle. Indicated are the relative input (RI) of the two excited states, the rate constants, the relative yield (RY) for I_0 formation from each of the excited states, and the quantum yield (QY) for I_1 formation obtained from fitting this model to the data.**

is present, therefore no attempt was made to extract any kinetic component in the order of the instrument response, i.e. ≤ 200 fs. The state associated with the 0.9 – 1 ns lifetime and the non–decaying state can be identified as the I_0 and I_1 states, respectively. The spectra of the 2 and 9 ps components are due to the excited state (ES) of PYP, which is known to decay multi–exponentially (*6–11*). We performed a target analysis (*12,13*) of the data, using a specific kinetic model based on results from visible femtosecond pump–probe experiments (*10–11*). This allows us to extract the species–associated difference spectra (SADS). The spectra of the two ES were virtually identical, which is why in a second round the data was fitted to a model in which the initial decay is bi–exponential (see Figure 3). The relative amplitudes of the 2ps and 9ps time constants were determined to be 0.7 and 0.3, respectively. The overall quantum yield of I_0 (and I_1) formation was estimated in the analysis to be 0.24 ± 0.05, of which the major part is formed with the 2 ps time constant, in good agreement with the yield of I_1 formation determined from visible data (*10–11*). The SADS of the states are shown in Figure 4. Note that in these spectra negative bands are due to the disappearance of *trans*–absorption, and positive (going) bands originate from excited– or product–state absorption.

The assignment of the dynamic bandshifts for each of the ES, I_0 and I_1 states is based on literature reports (*17, 18, 20, 21, 25–31*) and on experiments performed on a mutant in which the residue Glu46 (glutamate) has been replaced by a glutamine (see chapter 4), and is given in Table 1. The I_1 spectrum shows a very good resemblance to the ~50 ns – 1 µs FTIR difference spectrum reported by Brudler et al (*17*) and to the FTIR difference spectrum reported for I_1 in a cryotrapped intermediate (*18*). A difference with the spectrum reported by Brudler et al (*17*) is the amide I signal at –1643/+1626 cm^{-1} in their spectrum due to the response of the protein to the isomerization. On the time scale of our investigation such signals are absent; the spectra correspond quite well with

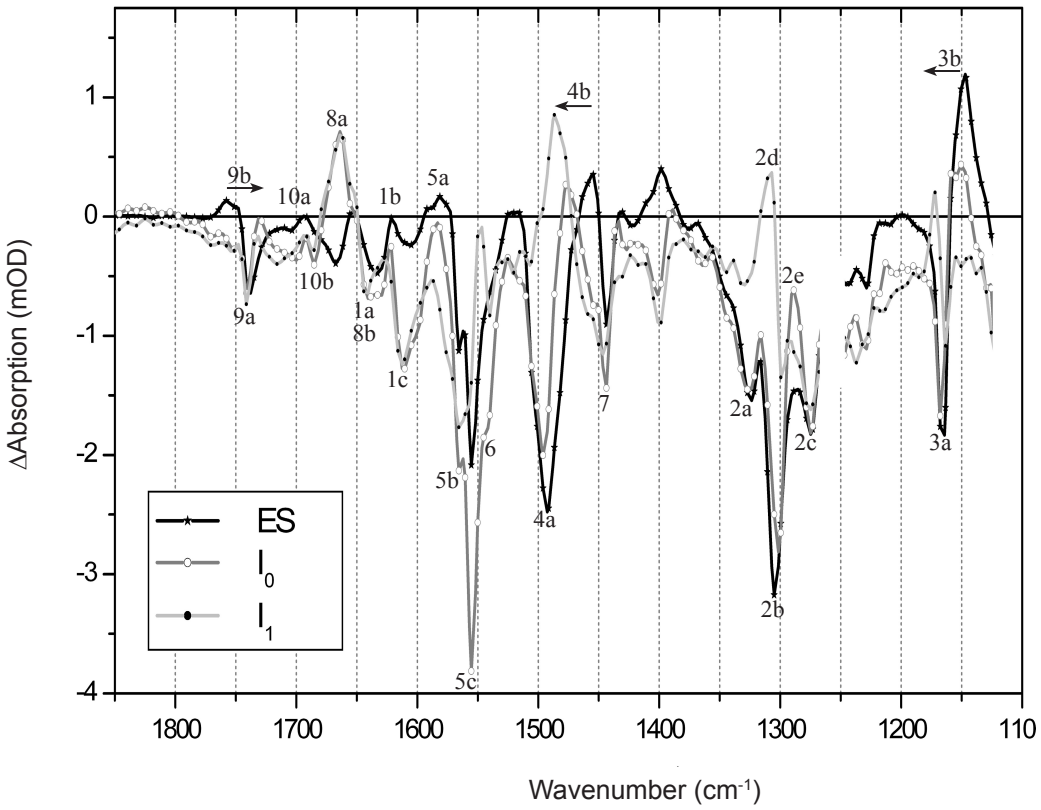


Figure 4 **Species–associated difference spectra of WT PYP induced by excitation of the chromophore with a short 475–nm laser pulse. The ES decay was fitted with a biexponential rate of (2 and 9 ps)$^{-1}$, the I_0 to I_1 transition rate was (0.9 –1 ns)$^{-1}$. Since the quantum yield of the I_0 and I_1 states was only 24%, their presence in the raw data was ~4 times smaller relative to the ES spectrum.**

Raman spectra *(25, 28,29*) and those of model chromophores in solution (*26,27*) and therefore can be assigned to structural changes in the chromophore itself, with the exception of bands 9 and possibly 10. Apparently the protein starts to respond on the 10–50 ns time scale. FTIR difference spectra have also been reported for the cryotrapped intermediates PYP_B and PYP_H, which, after forming PYP_{BL} and PYP_{HL}, respectively, convert to I_1. Note that there is a difference in visible absorption between I_0, PYP_B and PYP_H, (495, 490 and 440 nm, respectively). Nevertheless, our I_0 spectrum resembles both cryo–trapped intermediates quite well, but differs in the downshift of mode 3 while only PYP_B has a mode at 1485 cm^{-1}, close to our mode 4. Interestingly, these modes are both due to the phenol ring of the chromophore. We suggest that these differences are due to processes involving the chromophore or surrounding residues that are hampered at low temperature (see discussion below).

From the strong similarity of our transient spectra with published *trans–cis* difference spectra of model chromophores, the PYP_B, PYP_H and the I_1 state, we conclude that the pCA chromophore in both the I_0 and the I_1 transient state has isomerized.

The *trans* features in region 2 (ascribed to normal modes containing the C–C single stretches, including the C–C(–S–)=O

Frequency/Assignment		ES	I_0	I_1
-1633/+1621/-1607 C=C pCA *trans-cis* markers[17,18,26-29]	1a-c	~	+	+
-1326/-1302/-1274/+1308/+1289 C–C(–S–)=O pCA *trans-cis* markers[17,18,27,29]	2a-e	+	+	+
-1168 Phenol ring Y9a pCA *trans-cis* and ionic markers[18,27-29]	3a,b	+ 1149 b	+ 1155 b	+ 1173 b
-1495 Phenol ring Y19a, ring-O^- symmetric stretch, ionic marker[28,29]	4a,b	+ 1457 b	+ 1476 b	+ 1486 b
+(1580-1585)/-1565/-1555 Phenol ring Y8b ionic marker[25,27-30]	5a-c	+	+	–
-1540 Phenol ring Y8a[28,30]	6	+	+	+
-1444 Phenol ring Y19b[28,29]	7	+	+	+
+1665/-(1635-1640) C=O of pCA[25]	8a,b	–	+	+
-1740 COOH stretch of Glu46[18,20]	9a,b	+ 1755	+ 1732	+ 1732
+1693/-1685 Arg52 C=N[31]	10a,b	~+	+	+

Table 1 **Assignment of features observed in the midIR difference spectra for the states ES, I_0 and I_1. Frequencies are in cm^{-1}, in the left column +/– denotes a positive and negative band respectively, and in the right three columns +/~/– indicates the presence/absence of these bands in the different states. If present and different for each of the states, the frequency of the product band is indicated in the right three columns.**

stretch at 1302 cm^{-1} (*18,29*)) have disappeared already in the initial ES spectrum, and partly also those of the C=C stretch vibrations (reported at –1663/+1621/–1607 cm^{-1} for pCA methyl ester (*26*)). This may for a large part be explained by a decrease of pure single and double bond character, due to displacement of electronic charge from the C_7=C_8 isomerizable double bond into the adjacent single bonds upon excitation (*36*). Other changes observed in the ES spectrum are shifts of phenolate bands 3, 4 (ring–O^- symmetric stretch, refs *28,29*) and 5. At first sight these shifts seem to indicate that protonation changes of the phenolate ring occur: The upshift of the ~1560 cm^{-1} bands, the upshift of the 1168 cm^{-1} band in 850–950 ps, and the disappearance of the 1444 and 1495 cm^{-1} bands (*29*) are quite similar to the FTIR spectrum reported for PYP_M (37), in which the chromophore is protonated. It

is however unlikely that the chromophore actually gets protonated, since the 1755/1740/1732 cm^{-1} feature ('9') shows that the hydrogen bond with Glu46 is weakened, but not lost, in the ES and strengthened in I_0 and I_1, see below. Furthermore, protonation leads to a strong blue shift of the electronic absorption of the chromophore and this occurs on the sub–millisecond time scale. Recently, on the basis of time dependent DFT calculation it was suggested (*38*) that in the I_0 state (i.e. the cryotrapped structure reported by Genick et al (*19*) was used in these calculations) a subpopulation of the chromophore is protonated, which led to a much smaller electronic blue shift than that of the PYP_M state in line with visible pump–probe experiments. These authors found that in the I_1 state the chromophore was deprotonated again. Our dynamic shifts of the phenol bands might therefore be explained by a partial localization of the proton on the chromophore in the ES and I_0 states. Alternatively, or in addition, these modes may be sensitive to the chromophore becoming more neutral, due to the instantaneous change in dipole moment of 26 Debye upon light absorption, recently measured for PYP. This dipole moment change was interpreted as a charge translocation from the phenolic oxygen towards the ethylene chain, probably leading to a molecule with almost zero permanent dipole moment (*39*). Such a charge translocation to the ethylene chain may lead to a further decrease in pure C=C double and C–C single bond character (*39*), and therefore be important in initiating the *trans–cis* isomerization. For bacteriorhodopsin the presence of substantial light–induced charge redistribution in the retinal chromophore was shown to be a prerequisite for the occurrence of a photocycle (*40* and refs therein), therefore this motive seems to be a common trait in biological signal generation.

The C=O stretch of Glu46 at 1740 cm^{-1} upshifts to 1755 cm^{-1} in the ES spectrum, reflecting a weakening of the H–bond of Glu46 with the phenolic oxygen. This could be caused by a small increase in distance between Glu46 and the chromophore. However, in light of the observations discussed above, either the disappearance of the electron from the phenolic oxygen or the partial localization of the proton closer to the chromophore may be a more likely cause. Indeed, the upshift disappears with the ES to I_0 transition, when, as argued above, a more negative phenolic oxygen is reformed. In the I_0 and I_1 spectra subsequently a downshift of the Glu46 C=O is observed, indicating a *strengthening* of the Glu46–chromophore H–bond relative to the ground state structure, probably due to a small decrease in the Glu46 – chromophore distance. Probably, this initiates the proton transfer from Glu46 to the chromophore on the sub–ms time scale, leading to the swinging of the phenol ring, out of the protein pocket.

The frequency up–shift of the chromophore's C=O stretch from ~1640 to 1663 cm^{-1} (feature 8a,b) in the ES–to–I_0 transition clearly

demonstrates the breaking of the hydrogen bond with Cys69 (*25*), probably due to the flipping of the C=O around the ethylene chain of the chromophore. This flipping of the C=O group has been observed during the first few ns of the I_1 intermediate in time–resolved X–ray data (*16*) and in an early cryotrapped (PYP_{BL}) intermediate (*19*). Furthermore, an increase of bleaching of *trans* C=C modes at 1607 cm^{-1} is observed, and the appearance of an upgoing C–C product band at 1289 cm^{-1} typical for the formation of the *cis* isomer (*18*). The complete disappearance of the C=C (and C–C) *trans* modes, the appearance of C–C *cis* mode, together with the decrease in distance between Glu46 and the chromophore, and the breaking of the H–bond with Cys69, shows that the *cis*–ground state is formed on the picosecond time scale. Probably, the chromophore assumes in I_0 a stretched *cis*–configuration, similar to that observed in the cryotrapped PYP_B structure.

The changes in the region of the C–C(–S–)=O vibrational modes ('feature 2'), in particular the rise of the positive band at ~1310 cm^{-1}, show that a further *cis*–relaxation around these bonds takes place on the time scale of the I_0 to I_1 transition. The quantum yield for the I_0 to I_1 transition is estimated to be 90–100%, indicating that I_0 is a stable *cis* intermediate from which no return to the *trans* ground state occurs. Time resolved X–ray studies show that in the I_1 intermediate the chromophore is in a slightly stretched *cis*–conformation. In the first few ns the carbonyl oxygen is in a flipped position on the opposite side of the chromophore (*16*, see structure depicted in Figure 1B).

On no time scale breaking of the hydrogen bond with Glu46 is observed in our spectra, excluding the possibility of a large movement of the phenol ring of the chromophore during isomerization. Therefore isomerization does seem to occur via a double isomerization mechanism about the vinyl bond and the thioesther linkage in which the connections of the chromophore with the protein remain intact (*20,21*).

The small +1693/–1685 cm^{-1} band shift ('10') observed in all spectra we tentatively assign to Arg52, of which the C=N stretch lies in this frequency region (*31*). Groenhof *et al* (*32*) found in molecular dynamics simulations that upon excitation partial charge transfer from the chromophore to Arg52 takes place. These changes of Arg52 are in line with the time–resolved X–ray data of Ren *et al*. who observe, within a few ns, significant changes in residues not in direct contact with the chromophore (*16*). Here we show that these changes, probably due to the changed electron distribution/ polarization of the chromophore, are sensed within ~200 fs, and remain present during the ns time scale of the experiment.

Conclusions

Using time–resolved mid–IR spectroscopy, we have resolved

the structural changes in the very early part of the PYP photocycle. Our measurements are the first to characterize the vibrational features of the excited state from which isomerization occurs with a 20–30% overall quantum yield. We have found additional evidence for a charge translocation upon excitation from the phenolic oxygen towards the ethylene chain, which may be important for the weakening of the isomerizable C_7=C_8 double bond. Isomerization occurs on the 2–picosecond time scale and is accompanied by breaking of the hydrogen bond of the carbonyl oxygen with the protein. The transition from I_0 to I_1 occurs with a yield of 90–100% and we therefore conclude that I_0 is a stable long–living *cis* ground state configuration, which structurally relaxes on the nanosecond time scale of the I_0 to I_1 transition. We have further shown that the H–bond of the chromophore with Glu46 remains intact on all time scales relevant for this investigation. The combination of charge translocation and isomerization upon light absorption seems to be a common theme in photosensors. The protein appears to play an active role in combining the two, and thereby in directing the photocycle, since it stabilizes the negative charge on the phenolic oxygen in the ground state by an extensive hydrogen bond network.

Acknowledgements

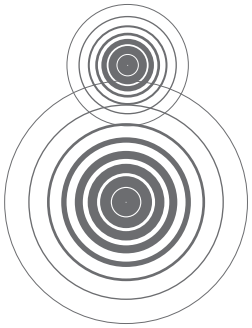
Assistance and technical support from Han Voet, Jos Thieme and the mechanical and electronic VU workshops is gratefully acknowledged.

References

1. Hellingwerf, K.J., Hendriks, J., Gensch, T. (2003) *J. Phys. Chem. A. 107*, 1082–1094
2. Meyer, T.E., Yakali, E., Cusanovich, M.A., Tollin, G. (1987) *Biochemistry 26*, 418–423
3. Hoff, W.D., Dux, P., Hard, K., Devreese, B., Nugteren–Roodzant, I.M., Crielaard, W., Boelens, R., Kaptein, R., Van Beeumen, J., Hellingwerf, K.J. (1994), *Biochemistry 33*, 13959–13965
4. Baca, M., Borgstahl, G.E.O., Boissinot, M., Burke, P.M., Williams, D.R., Slater, K.A., Getzoff, E.D. (1994) *Biochemistry 33*, 14369–14377
5. Van Brederode, M.E., Hoff, W.D., van Stokkum, I.H.M., Groot, M.L., Hellingwerf, K.J. (1996) *Biophys. J. 71*, 365–380
6. Baltuška, A., van Stokkum, I.H.M., Kroon, A., Monshouwe, R., Hellingwerf, K.J., van Grondelle, R. (1997) *Chem. Phys. Lett. 270*, 263–266
7. Ujj, L., Devanathan, S., Meyer, T.E., Cusanovich, M.A., Tollin, G., Atkinson, G.H. (1998), *Biophys. J. 75*, 406–412
8. Devanathan, S., Pacheco, A., Ujj, L., Cusanovich, M.A., Tollin, G., Lin, S., Woodbury, N., (1999) *Biophys. J. 77*, 1017–1023
9. Imamoto, Y., Kataoka, M., Tokunaga, F., Asahi, T., Masuhara, H. (2001) *Biochemistry 40*, 6047–6052
10. Gensch, T., Gradinaru, C.C., van Stokkum, I.H.M., Hendriks, J., Hellingwerf, K.J., van Grondelle, R. (2002) *Chem. Phys. Lett. 356*, 347–354
11. Larsen, D.S., Vengris, M., van Stokkum, I.H.M., van der Horst, M.A., de Weerd, F.L., Hellingwerf, K.J., van Grondelle, R. submitted for publication
12. Holzwarth, A.R. (1996) "Data analysis of time–resolved measurements," in *Biophysical Techniques in Photosynthesis* (Eds. Amesz, J. and Hoff, A. J.) pp. 75–92, Kluwer, Dordrecht
13. Hoff, W.D., van Stokkum, I.H.M., van Ramesdonk, H.J., van Brederode, M.E., Brouwer, A.M., Fitch, J.C., Meyer, T.E., van Grondelle, R., Hellingwerf, K.J. (1994) *Biophys. J. 67*, 1691–1705
14. Hendriks, J., van Stokkum, I.H.M., Hellingwerf , K.J. (2003) *Biophys. J. 84*, 1180–1191
15. Perman, B., Srajer, V., Ren, Z., Teng, T., Pradervand, C., Ursby, T., Bourgeois, D., Schotte, F., Wulff, M., Kort, R., Hellingwerf, K.J., Moffat, K., (1998) *Science 279*, 1946–1950
16. Ren, Z., Perman, B., Srajer, V., Teng, T.–Y., Pradervand, C., Bourgeois, D., Schotte, F., Ursby, Th., Kort, R., Wulff, M., Moffat, K. (2001) *Biochemistry 40*, 13788–13801
17. Brudler, R., Rammelsberg, R., Woo, T.T., Getzoff, E.D., Gerwert, K. (2001) *Nat. Struct. Biol. 8*, 265–270
18. Imamoto, Y., Shirahige, Y., Tokunaga, F., Kinoshita, T.,

Yoshihara, K., Kataoka, M. (2001) *Biochemistry 40*, 8997–9004
19. Genick, U.K., Soltis, S.M., Kuhn, P., Canestrelli, I.L., Getzoff, E.D. (1998) *Nature 392*, 206–209
20. Xie, A., Hoff, W.D., Kroon, A.R., Hellingwerf, K.J. (1996) *Biochemistry 35*, 14671–14678
21. Imamoto, Y., Kataoka, M., Liu, R.S.H. (2002) *Photochem. Photobiol. 76*, 584–589
22. Groot, M.L., Vos, M.H., Schlichting, I., Mourik, F. van, Joffre, M., Lambry, J.–C., Martin, J–L. (2002) *Proc. Natl. Acad. Sci. USA 99*,1323–1328
23. Hendriks, J., Gensch, T., Hviid, L., van der Horst, M.A., Hellingwerf, K.J., van Thor, J.J. (2002) *Biophys. J. 82*, 1632–1643
24. Genick, U.K., Borgstahl, G.E.O., Ng, K., Ren, Z., Pradervand, C., Burke, P.M., Srajer, V., Teng, T–Y., Schildkamp, W., McRee, D.E., Moffat, K., Getzoff, E.D. (1997) *Science 275*, 1471–1475
25. Unno, M., Kumauchi, M., Sasaki, J., Tokunaga, F., Yamauchi, S. (2002) *Biochemistry 41* 5668–5674
26. Xie, A., Kelemen, L., Hendriks, J., White, B.J., Hellingwerf, K.J., Hoff, W.D. (2001) *Biochemistry 40*, 1510–1517
27. Van Thor, J.J., Pierik, A.J., Nugteren–Roodzant, I., Xie, A., Hellingwerf, K.J. (1998) *Biochemistry 37*, 16915–16921
28. Zhou, Y., Ujj, L., Meyer, T.E., Cusanovich, M.A., Atkinson, G.H. (2001) *J. Phys. Chem. A 105*, 5719–5726
29. Kim, M., Mathies, R.A., Hoff, W.D., Hellingwerf, K.J. (1995) *Biochemistry 34*, 12669–12672
30. Brudler, R., Meyer, T.E., Genick, U.K., Devanathan, S., Woo, T.T., Millar, D.P., Gerwert, K., Cusanovich, M.A., Tollin, G., Getzoff, E.D. (2000) *Biochemistry 39*, 13478–13486
31. Tamm, L.K., Tatulian, S.A. (1997) *Quarterly Reviews of Biophysics 30*, 365–429
32. Groenhof, G., Lansink, M.F., Berendsen, H.J.C., Snijders, J.G., Mark, A.E. (2002) *Proteins: Structure, Function and Genetics 48*, 202–211
33. Gensch, T., Hellingwerf, K.J., Braslavsky, S.E., Schaffner, K. (1998) *J. Phys. Chem. A 102*, 5398–5405
34. Herbst, J., Heyne, K., Diller, R. (2002) *Science 297*, 822–825
35. Yamada, A., Yamamoto, S., Yamato, T., Kakitani, T. (2001) *J. Mol. Struct. (Theochem) 536*, 195–201
36. Sergi, A., Gruning, M., Ferrario, M., Buda, F. (2001) *J. Phys. Chem. B 105*, 4386–4391
37. Imamoto, Y., Mihara, K., Hisatomi, O., Kataoka, M., Tokunaga, F., Bojkova, N., Yoshihara, K. (1997) *J. Biol. Chem. 272*, 12905–129808
38. Thompson, M.J., Bashford, D., Noddleman, L., Getzoff, E.D.

(2003) J. Am. Chem. Soc. In press
39. Premvardhan, L.L., van der Horst, M.A., Hellingwerf, K.J., van Grondelle, R.(2003) *Biophys J. 84*, 3226–3239
40. Zadok, U., Khatchatouriants, A., Lewis, A., Ottolenghi, M., Sheves, M. (2002) *J. Am. Chem. Soc. 124*, 11844–11845



Ultrafast Infrared Spectroscopy reveals Key Step for Successful Entry into Photocycle for Photoactive Yellow Protein

04

Based on: L.J.G.W. van Wilderen, M.A. van der Horst, I.H.M. Van Stokkum, K.J. Hellingwerf, R. van Grondelle, M.L. Groot Proceeding of the National Academy of Sciences 2006 103 (41) 15050-15055 www.pnas.org/cgi/doi/10.1073/pnas.0603476103

Abstract

Photoactive proteins like the Photoactive Yellow Protein (PYP) are generally accepted as model systems for studying protein signal state formation. PYP is a blue light sensor from the bacterium Halorhodospira halophila. The formation of PYP's signalling state is initiated by trans-cis isomerization of the p-coumaric acid (pCA) chromophore upon the absorption of light. The quantum yield of signalling state formation is approximately 0.3. Using femtosecond visible pump/mid-IR probe spectroscopy, we investigated the structure of the very short-lived ground state intermediate that results from an unsuccessful attempt to enter the photocycle. This intermediate and the first stable ground state intermediate on-pathway in the photocycle, I_0, both have a mid-IR difference spectrum characteristic of a cis*-isomer, but only the I_0 intermediate has a chromophore with a broken hydrogen bond with the backbone N-atom of Cys69. Therefore, we suggest that breaking of this hydrogen bond is decisive for a successful entry into the photocycle. The chromophore also engages in a hydrogen-bonding network via its phenolate group with residues Tyr42 and Glu46. We have investigated the role of this hydrogen bond by exchanging the H-bond donating residue Glu46 with the weaker H-bond donating glutamine (i.e. Gln46). We have observed that this mutant exhibits virtually identical kinetics and product yields as wild type PYP, even though during the I_0 to I_1 transition, on the 800-ps time scale, the hydrogen bond of the chromophore with Gln46 is broken, whereas this hydrogen bond remains intact with Glu46.*

Introduction

PYP belongs to the Xanthopsins, a family of blue-light photoreceptors which contain 4-hydroxy-cinnamic acid as their photoactive chromophore (see (1-3) for a review). PYP is a small protein and therefore an attractive model system to explore how a chromophore and the protein interact to sense light and send a biological signal. Its photocycle has been characterized by various experimental techniques, such as fluorescence (4, 5), (time-resolved) FTIR (6-8), (time-resolved) X-ray crystallography (9-12), NMR (13), Stark spectroscopy (14) and pump(-dump)-probe spectroscopy (15, 16). X-ray diffraction on PYP-crystals has demonstrated that the PYP-chromophore is covalently linked (see Figure 1) to the protein backbone via Cys69 (12). It is further embedded in a hydrogen bonding network, consisting of Glu46, Tyr42, Thr50 and Cys69 (12). In the ground state, the chromophore is in a deprotonated *trans*-form, negatively charged, and possibly stabilized by the positive Arg52 residue (12). After photoexcitation and within a few picoseconds, the chromophore forms a red-shifted intermediate, referred to as I_0. This intermediate has a shifted absorption maximum from 446 to 500 nm. The second intermediate I_1 (or pR or PYP_L), absorbs maximally at 480 nm, and is formed in 1-3 ns (15-20). This is followed by protonation of the chromophore and a large structural change of the protein on a millisecond time scale, which is believed to be the signalling state (see (1-3) for a review). A previous report of visible pump/mid-IR probe measurements on wild type (WT) Photoactive Yellow Protein ((21) and chapter 3) provided detailed insight in the initial structural changes taking place during chromophore isomerization on the 200 fs to 3 ns time scale, a period covering formation of the excited state (ES), I_0 and I_1. A stable *cis*-ground state was observed to form in about 2 ps which was confirmed by Heyne and co-workers using the same technique, in combination with normal mode calculations (22). In conjunction with isomerization of the chromophore, dynamic changes of the hydrogen bonding network surrounding the chromophore were observed. For example, the carbonyl group of the chromophore breaks its hydrogen bond to the back bone of Cys69 (21) and probably flips to the other side of the pocket, leading to the conformation observed for a cryotrapped intermediate in X-ray diffraction and during the first few nanoseconds of the I_1 state using time-resolved X-ray diffraction experiments (9, 23). The hydrogen bond between the chromophore and Glu46 is weakened during the initial events of the photocycle and subsequently strengthened (21).

In this study, we investigate the functional role of the hydrogen bonding network of the PYP chromophore by comparing the mid-IR difference spectra of the PYP photocycle intermediates in the Glu46Gln (shorthand notation: E46Q) mutant with those of WT PYP. Glutamine donates a weaker hydrogen bond to the

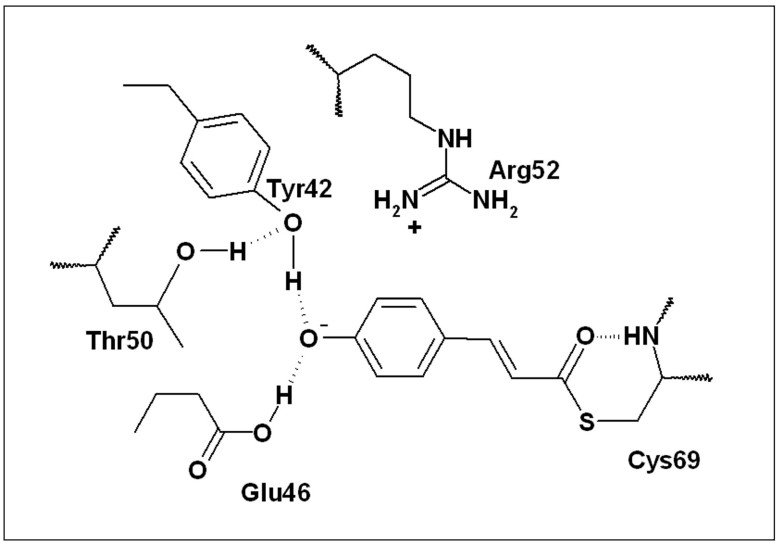


Figure 1 **Schematical drawing of the active site of WT PYP. The *p*-coumaric acid chromophore is covalently bound to the protein backbone via Cys69; in addition, it takes part in a distal hydrogen bonding network. In E46Q Glutamic acid (Glu) is changed to Glutamine (Gln) using site-directed mutagenesis.**

chromophore because of the inherently weaker hydrogen bonding characteristics of the amide group of the glutamine side chain (24, 25), but in the ground state the hydrogen bond is still present (26, 27). This mutant has a considerably red-shifted absorption spectrum compared to WT PYP; it peaks at 460 nm. Its photocycle is about 3 times faster (28) (recovery lifetime of 50 ms instead of 140 ms, but this is strongly pH-dependent).

In addition we focus on the structural events during the isomerization process which are responsible for successful entrance into the photocycle. Recently, using visible pump-dump probe spectroscopy, an early (ground state) intermediate other than I_0 was identified (15). This intermediate was observed to form from the excited state in competition with I_0, but decayed to the original ground state in about 3-4 ps. This intermediate therefore is originating from unsuccessful attempts of the chromophore to enter the photocycle. The quantum yield for successful entry into the photocycle is ~ 0.3 (15, 21, 29). Comparing the mid-IR difference spectra of the two intermediates might yield information on which molecular factors are responsible for the relatively low yield of signaling state formation. We build upon the results of the aforementioned pump-dump-probe experiments and specifically introduce this ground state intermediate in our data analysis. The vibrational absorption-difference spectrum of this state reveals that the chromophore is structurally distorted and can most likely be considered to be a *cis*-isomer but with the hydrogen bond between the chromophore's C=O group and the Cys69 residue of the protein still intact.

Materials and Methods

The sample was prepared in H_2O buffer, as previously described

(45). It consisted of a highly concentrated protein solution pressed between two CaF_2 windows, separated by a 20 μm Teflon spacer, at an OD of about 1.0 at 446 nm. The experimental setup (see (21) for a more detailed description) consisted of an integrated Ti:sapphire oscillator/ regenerative amplifier, operating at 1 kHz, and producing 0.8 mJ pulses of 80 fs (Hurricane, SpectraPhysics). The output of this laser was used to pump a commercial optical parametric generator and amplifier with difference frequency generation (TOPAS, Light Conversion), which resulted in a tuneable output (2.5-10 μm) with a spectral width of ~200 cm^{-1}. A home-built HgCdT camera system placed behind a spectrograph was read out every shot at a repetition rate of 1 kHz and a sampling resolution of ~6 cm^{-1}. Another part of the Hurricane output was used to pump a home-built non-collinear OPA, to generate pulses at 475 nm, with duration of ~60 fs (uncompressed) and 90 nJ excitation energy. Pump-probe spectra were measured between –10 ps and 3 ns with an instrument response function of about 200 fs (cross-correlation in GaAs). The pump beam polarization was set with a Berek-rotator to magic angle with respect to the probe beam. A phase-locked chopper at 500 Hz ensured that every other shot the sample is excited and an absorbance difference spectrum could be calculated. To ensure a fresh spot for each laser shot, the sample was moved with a homebuilt Lissajous scanner. The setup was enclosed in a nitrogen-purged box to reduce distortions of the infrared beam by water vapour absorption.

Data was collected in the spectral window from 1090 to 1740 cm^{-1} in 5 partially overlapping windows. To check reproducibility, every dataset, typically consisting of 60 scans taking about 70 minutes in total, was measured several times. The data was processed and analysed by global and target analysis software (30). A correction was made for the presence of a pre-time-zero offset due to thermal lens effects (46, 47). Since no reference probe pulse was used, the noise in the measured spectra consisted mainly of baseline noise, i.e. a flat, structureless offset in the spectra, which was easily recognized from a singular value decomposition of the residual matrix (30, 48). The signal to noise ratio of the data was enhanced by subtracting the outer product of the first two singular value pairs of the residual matrix (being structureless in the time domain and smooth in the wavelength domain) from the data, leading to a factor 2 reduction in the noise for both WT and E46Q. The previously recorded WT data (21) is re-analysed with these two corrections. A good quality data set had a typical noise level smaller than 100 μOD.

Results

The PYP E46Q absorption-difference data, photo-excited at 475 nm, consist of 160 timetraces recorded between 1090 and 1740 cm^{-1} (see Figure 2). A global analysis, i.e. of all time traces

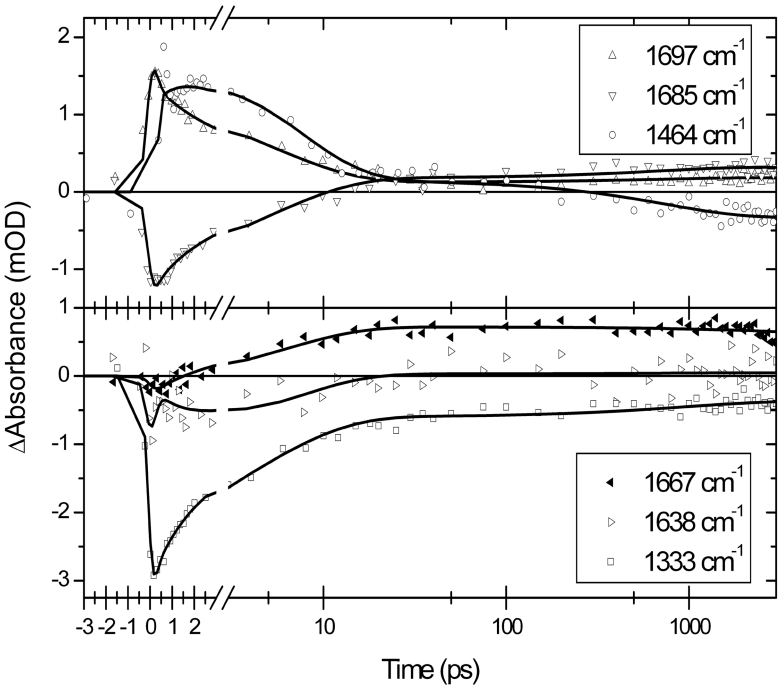


Figure 2 **Selection of time traces measured on E46Q PYP. Absorption difference (mOD) is plotted as a function of time (ps). The time-axis is linear up to 3 ps and logarithmic until 3 ns. The solid lines are the fit through the data, based on a target analysis.**

simultaneously, with a model of parallel decaying states, showed that the data are well described using 4 exponential decays with time constants of 1.4 (± 0.1) ps, 6 (± 0.3) ps, 800 (± 70) ps and >10 ns, respectively. The corresponding decay-associated difference spectra (DADS) of each of these time constants in the 1740-1610 cm^{-1} region are shown in Figure 3. Because of the presence of a small perturbed free induction decay before *t = 0* no information faster than the instrument response (~200 fs) is extracted. The two longest time constants can be identified as the lifetimes of the states I_0^{E46Q} and I_1^{E46Q}, respectively. The 1.4 ps time constant is similar to the value(s) reported for the excited state lifetime in WT PYP; a multi-exponential decay with 0.6 ps and 2.8 ps lifetimes was observed using visible pump-probe spectroscopy (15) and 2 ps and 3 ps time constants were reported using vis/midIR pump-probe spectroscopy (21, 22)*. Therefore we assign the 1.4 ps time constant to the excited state lifetime of E46Q. The 6 ps component is associated with spectral changes which are distinctly different from those of the 1.4–ps excited state component and they also clearly differ from the spectral characteristics of the I_0^{E46Q}-state

* The inclusion of a fast, 0.7 ps (fixed) time constant with amplitude 10% relative to that of the 1.4 ps time constant, led to a small improvement of the fit. However, since we could not resolve it independently from our data and since it possessed a spectrum identical to that of the 1.4 ps component, we ignored its presence. Apparently the vis-vis pump probe experiments are more sensitive to this fast initial relaxation of the excited state than the midIR experiments, indicating that this relaxation is most likely solvational rather than structural.

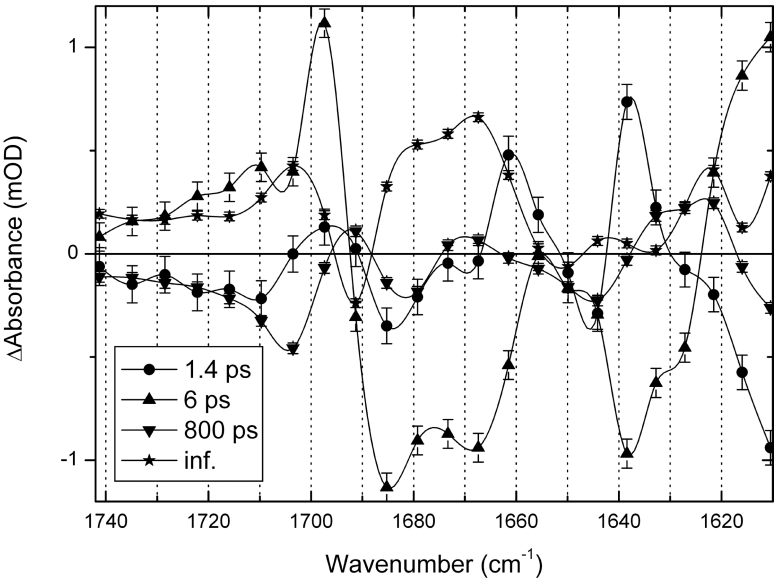


Figure 3 **Decay associated difference spectra (including their error bars) of the E46Q data, fitted with four time constants.**

(see Figure 3).

The initial dynamics of PYP have been shown to involve the parallel decay of the excited state into the I_0 product state and into a short-lived ground state intermediate (15). To extract the spectra of these states and to estimate their relative quantum yields, we have analysed our data using a target model (30) based on that reported by Larsen et al (15). In this model, the excited state can decay into the I_0 ground state and in parallel into a short-lived ground state intermediate denoted GSI (see Figure 4). From I_0 the I_1 state is formed, whereas GSI decays into the stable ground state. When this model is applied to the data we find that the 1.4 ps, 6 ps, 800 ps and 10 ns time constants correspond to the lifetimes of the excited state, GSI^{E46Q}, I_0^{E46Q} and I_1^{E46Q}, respectively. The branching fractions, giving the relative contribution of the decay from the excited state into I_0^{E46Q} and GSI^{E46Q}, equal 0.31 ± 0.05 and 0.58 ± 0.05, respectively (see Table 1). The formation of I_0^{E46Q} seems to be the point of no return, as I_1^{E46Q} is formed from I_0^{E46Q} with a yield of 90-100%. The overall quantum yield for I_1^{E46Q} formation, 0.27-0.31, is similar to that reported before for WT PYP; 0.35-0.5 (18, 21, 30). Alternative models were explored to test the possibility of a vibrational cooling process (i.e. by introducing a state in between ES and I_0) and the formation of a completely separate state. Based on the resulting spectra and quantum yield, as determined for several long-living signals, we conclude that our data are best described by the presented model where two states are formed in parallel from the excited state, and the second species (GSI) falls directly back into the equilibrated groundstate, consistent with Larsen et al (15).

Because of the detection of the GSI intermediate in the IR data, we re-analysed the WT PYP data published earlier (21) with the

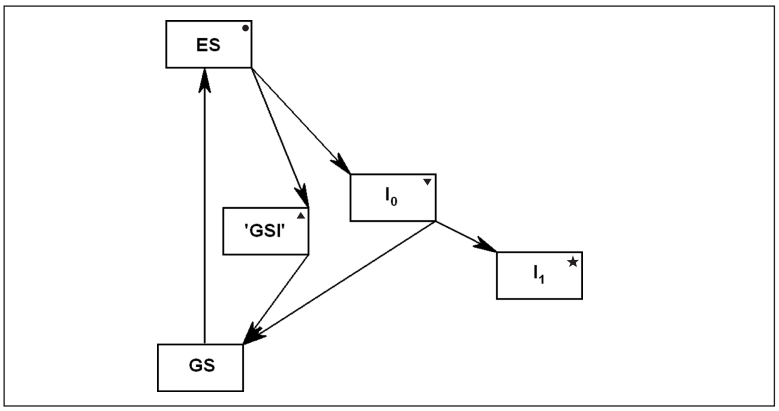


Figure 4 **Model used for the target analysis. ES is the excited state; GSI is a short-lived, ground state intermediate; GS represents the ground state; and I_0 and I_1 the first two transient, on-pathway, ground-state intermediates. The symbols correspond to the species they represent in Figure 3.**

added assumption of the presence of the GSI-state. The obtained results are shown in Table 1. The values are reasonably close to those reported by Larsen et al (the 'homogeneous model' in ref. (15)). The species associated difference spectra (SADS) of WT and E46Q that result from the target analysis are shown in Figure 5, and will be discussed in more detail below.

Discussion

The vibrational spectrum of a protein or a protein-bound chromophore contains a wealth of information about its structure, its interaction with the environment and its electronic properties. In addition, monitoring reaction-induced infrared absorption changes can reveal the response of those parts of the protein that are directly involved in the on-going reactions. The spectra that we have resolved for the initial events of the PYP photocycle yield information on crucial questions relating to protein-chromophore interactions and function. In particular, by comparing the spectra of the I_0 and the GSI states, we can characterize the factors that are important for a successful entry into the photocycle.

The spectra we obtained in this study were interpreted using previous assignments of FTIR and Raman spectra of model

Table 1 **Target analysis parameters**

	E46Q		WT	
State	Yield*	Time†	Yield*	Time†
ES	1.0	1.4 ps (0.1)	1.0	1.2 ps (0.1)
GSI	0.58‡	6.0 ps (0.3)	0.72	6.2 ps (0.3)
I_0	0.31‡	800 ps (70)	0.28	700 ps (50)
I_1	0.9-1.0§	>10 ns	0.9-1.0§	>10 ns

* Error in quantum yield is ± 0.05.
† Time constant; error is shown in parentheses.
‡ Remaining ES^{E46Q} (0.11) falls back into ground state.
§ Remaining I_0 falls back into ground state.

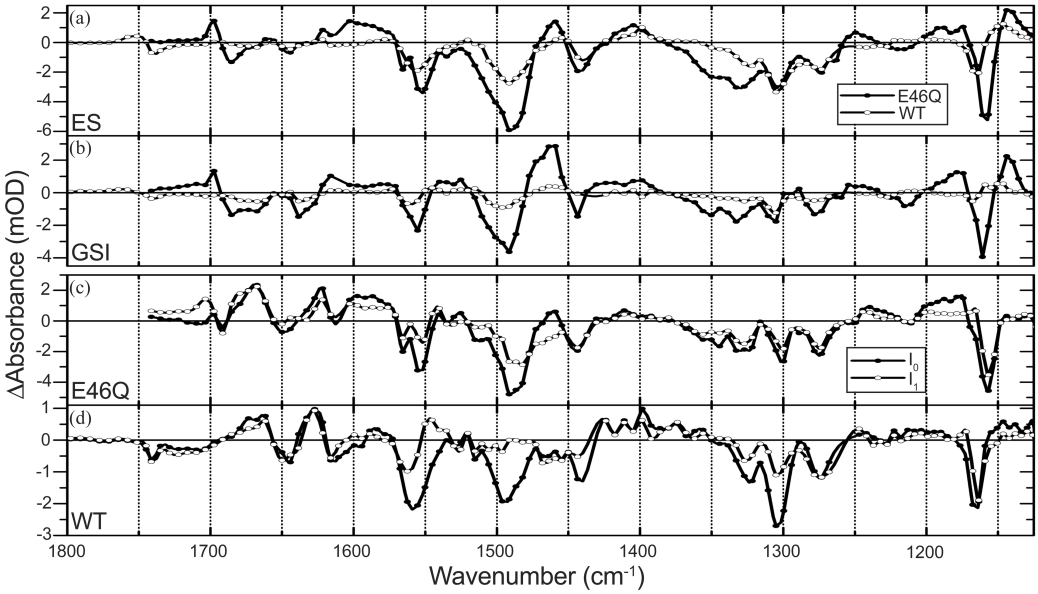


Figure 5 **SADS of WT and E46Q as a result of the target analysis. Comparison of the two excited state spectra of WT (open spheres) and E46Q (closed spheres) (a), and those of the ground state intermediates GSI^{WT} (open spheres) and GSI^{E46Q} (closed spheres) (b). The I_0 (closed spheres) and I_1 (open spheres) spectra for E46Q and WT PYP are shown in (c) and (d), respectively. Note that the bottom two panels share the same legend. Negative features in these spectra originate from the ground state bleach, and positive ones from ES- or product state(s) absorption.**

chromophores in solution, isotope labelled PYP, mutants, and normal mode analysis (7, 8, 22, 31-36). Note that the negative bands in the spectra correspond to those of PYP in its ground state whereas the positive bands arise from band shifts or product bands due to the formation of the ES, GSI, I_0 and I_1 states, respectively.

Isomerization of the chromophore in the I_0 and I_1 states

The phenolate-ring modes dominate the signals around 1160, 1443, 1485 and 1550 cm^{-1} (Figure 5). The instantaneous spectral changes of the phenolate-ring modes are similar to those of a protonated minus deprotonated *p*-coumaric acid difference spectrum (35) and imply that the phenol-ring becomes less negative in the excited state, since they all are also ionic markers. Electric field measurements on WT and E46Q PYP (14) showed an instantaneous change in dipole moment upon photoexcitation, which was explained by a charge translocation from the phenolic oxygen over the ethylene chain, in line with our observations. Note that this charge translocation would result in a rearrangement of the carbon bond conjugation of the chromophore, facilitating isomerization; It could give the double bond of the carbonyl group a more single bond character. Molecular dynamics simulations also indicated that the negative charge moves along the ethylene chain, although it was calculated to move all the way towards the nearby positively charged Arg52 residue (37), as opposed to movement towards the thio-ester linkage of the chromophore, as was deduced from the results of Stark spectroscopy (14).

Signals originating from C-C modes appear mainly in the 1280-1330 cm^{-1} region; the negative features at 1300/05 and 1327/33 cm^{-1} are C-C *trans-cis* markers (8, 32), indicating the

disappearance of the *trans* ground state of the chromophore. The upgoing band at 1289 cm^{-1} is a marker for the *cis*-product state. Structurally it has been assigned to the –C=C–C(–S–)=O skeleton stretch (31, 32).

The C=C vibrations (also *trans*-*cis* markers) appear mainly around 1600-1635 cm^{-1}, and those of the C=O of the chromophore in the 1640-1670 cm^{-1} region. This spectral region deserves special attention, and is discussed in more detail below. In general, however, the WT I_1 spectrum shows a very good resemblance with the steady-state FTIR (difference) spectrum of cryo-trapped I_1^{WT} (or PYP_L) (8) and with the time-resolved 50 ns step-scan FTIR spectrum reported by Brudler et al (6), as noted previously (21). I_1^{WT} has been demonstrated to be a *cis*-isomer in time-resolved X-ray Laue diffraction experiments (38). From the disappearance of the C-C and C=C *trans* modes, and the appearance of the C-C mode at 1289 cm^{-1}, we concluded that also in I_0^{WT} the chromophore had isomerized (21). This was confirmed by the observation of an additional *cis*-marker mode at 1000 cm^{-1} in the I_0^{WT} state (22). Changes in the C–C(–S)=O bond region occurring in ~800 ps (see features around 1300 cm^{-1}) indicate that a relaxation of the *cis*-configuration marks the I_0^{WT} to I_1^{WT} transition. The spectra of I_0 and I_1 of E46Q and WT PYP are overall very similar in these regions. We conclude therefore that in E46Q the chromophore follows the same initial events as in WT: the *cis*-isomer ground state I_0, is formed in 1.4 ps (slightly faster in WT), followed by a structural relaxation in 800 ps with the I_0 to I_1 transition.

Hydrogen bond breaking in E46Q

The E46Q spectra differ from WT around 1160 cm^{-1} and in the 1700-1740 cm^{-1} region. The negative band at 1165 cm^{-1} in WT PYP has shifted to 1160 cm^{-1} in the mutant. This band has been assigned to phenolate-ring mode Y9a (32-34) and is apparently sensitive to the altered hydrogen bond strength, due to replacement of Glu by Gln. The other main difference is in the C=O region of the (mutated) amino acid at position 46: For WT we observe band shifts in the 1735-1760 cm^{-1} region due to the C=O mode of Glu46, which are absent in the mutant. On the other hand, new signals appear in E46Q between 1685-1705 cm^{-1} that we assign to the C=O mode of Gln46 (see 'double' frame in Figure 6). This C=O mode shifts from 1685 cm^{-1} in the ground state to 1697 cm^{-1} in the excited- (spheres in Figure 6), GSI^{E46Q} (up-triangles) and I_0^{E46Q} states (down-triangles). In I_1^{E46Q} (stars) it further shifts to 1704 cm^{-1}, which is the same frequency as that reported for this state by Brudler et al using FTIR spectroscopy with nanosecond time resolution (6). With time-resolved X-ray Laue diffraction spectroscopy, it was shown that the hydrogen bond between Gln46 and the chromophore is lost in I_1 (38). We thus demonstrate here that disruption of this hydrogen bond occurs

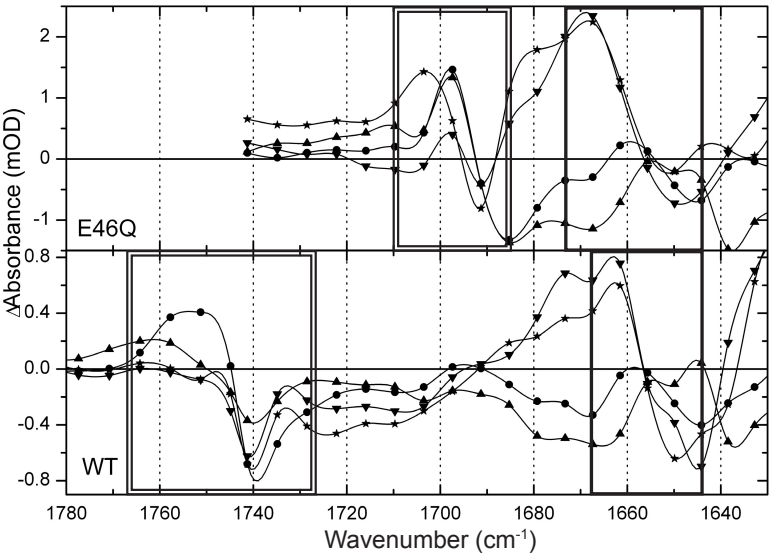


Figure 6 **Comparison of the SADS of E46Q (top) and WT (bottom) in the Glu/Gln C=O ('double' frame) and the chromophore C=O ('single' frame) regions. The spheres denote the ES, the up-triangles the GSI, the down-triangles I_0 and the stars I_1.**

during the I_0^{E46Q} to I_1^{E46Q} transition, on the 800-ps time scale. In WT PYP the hydrogen bond between Glu46 and the chromophore is still intact on this time scale, and is even slightly strengthened, see Figure 6 and (21).

The initial photocycle events, such as the rate and yield of I_0 and I_1 formation, are very similar for WT and E46Q, in spite of the weaker hydrogen bond in E46Q and its breakage on the 800-ps time scale. We conclude therefore that the strength of the hydrogen bond between the chromophore and residue 46 does not seem to play a crucial role in the initial part of the photocycle. At longer time scales, however, this mutation causes a faster ground state recovery rate (28, 39). Different ground state recovery pathways for WT and E46Q were also reported to show up in dehydrated films (40).

Successful or unsuccessful photocycle entry attempts

Larsen et al have shown that the state we resolve here as GSI, is a ground state intermediate through which a major part of the excited state decays to the ground state (15). The SADS of this state differs from both the I_0 and I_1 spectra, most notably in the 1600-1700 cm^{-1} region (see Figure 5).

In this region there is a positive band at 1667 cm^{-1} in the I_0^{E46Q} and I_1^{E46Q} spectra (see Figure 6, 'single' frame), which has been assigned to the chromophore's C=O group, free from interaction with the N-backbone atom of Cys69 (6, 36). Breaking the hydrogen bond, which occurs because of the rotation of the C=O group around the chromophore's long axis as observed in nanosecond time-resolved X-ray studies (38) and cryo-trapped intermediates (41), causes the frequency of this mode to upshift, leaving a bleached band at 1650 cm^{-1} in the I_0^{E46Q} and I_1^{E46Q} spectra. Remarkably, the positive features are notably absent in ES and GSI, indicating that

the hydrogen bond between the C=O of the chromophore and Cys69 is intact in ES and GSI.

We note that although there is no upshifted product band, the ES spectrum does contain a bleached band at 1644 cm^{-1}. This could be due to a more single bond character of the carbonyl group in the excited state, due to the electron migration upon excitation as discussed above. Such migration would shift the mode into the ~1200 cm^{-1} region, where it is unfortunately hard to recognize. In addition, there is a negative band in the GSI spectrum at 1638 cm^{-1}. This band has been proposed to originate from mainly the central C=C bond (32). The bleaching of this band could therefore be interpreted as that GSI assumes a backbone conformation, slightly different from I_0. We note that in the I_1 state as well as in the cryo-trapped PYP_L spectrum both the 1650 and the 1638 cm^{-1} band are fully resolved (8).

The spectral changes in the remaining part of the GSI spectrum, in particular in the 1400-1200 cm^{-1} region, which is dominated by C–C(–S–)=O modes, and the appearance of a positive *cis*-marker at 1289 cm^{-1}, show that the *trans* backbone of the chromophore in the GSI state has been distorted and could even be considered to be in a *cis* configuration, although, as noted above, the hydrogen bond of the C=O of the chromophore to Cys69 is still intact. Probably, the inability of the chromophore to break the hydrogen bond on a sufficiently fast time scale leads to its decay to the original ground state, via a mixture of vibrational cooling and re-isomerization to the *trans*-configuration, rather than to its entry into the photocycle by a relaxation of the *cis*-isomer. The important role of the strength of the chromophore's hydrogen bond with the backbone N-atom of Cys69 in determining the success of the isomerization is compatible with the *higher* quantum yield of isomerization in the P68A mutant where this hydrogen bond is weaker (42).

Isomerization mechanism

The observation of a short-lived ground state intermediate with the chromophore in a distorted *trans*/*cis* configuration and its hydrogen bond of the C=O group to backbone intact, is in agreement with the molecular dynamics simulations performed by Groenhof et al (43). They showed that breaking of the hydrogen bond is uncoupled from the isomerization process: In their simulations, isomerization around the C=C double bond takes places partly in the excited state, and continues rapidly with the reformation of the ground state. This is followed by breaking of the hydrogen bond (i.e. in the *cis*-ground state) in a few picoseconds (43), in agreement with our observation of a *cis*-isomer with its hydrogen bond intact. However, in these simulations an additional intermediate is predicted between ES and I_0, which has not been observed in our measurements. Possibly, our data does not contain

a sufficient signal to noise ratio to confirm this, or the concentration of an I_0-like intermediate with the C=O hydrogen bonded could be too low in the experiment, due to a faster decay- than formation rate. The large number of unsuccessful isomerization attempts observed in the MD simulations, could be consistent with the existence of a GSI. An initial inspection has revealed that these unsuccessful isomerizations originate from a 'twisted-state' formation (G. Groenhof, unpublished). The involvement of twisted charge-transfer states in the photocycle of PYP, either on- or off-pathway, deserves further attention. The unsuccessful attempts forming the GSI favour the hula-twist mechanism (see ref. (44)) since this mechanism is consistent with an intact hydrogen bonded C=O. For the entry into the photocycle, however, the multiple-bond flip mechanism apparently prevails since we do not observe an I_0 intermediate with the H-bond intact. However, as mentioned above, experimental difficulties could have hampered the observation of such a state, and consequently, this question is still open. Note that in Groenhof's simulations a successful entry into the photocycle occurs via the hula-twist mechanism.

Conclusions

Overview of the initial events of the PYP Photocycle

The envisioned sequence of events that occur in PYP upon excitation, based on our current results and the MD simulations by Groenhof et al (43) is the following: In the excited state, the negative charge on the phenolic oxygen migrates towards the thio-ester linkage, giving the double bond of the hydrogen bonded carbonyl group a more single bond character. The hydrogen bond to the backbone remains initially intact. Isomerization proceeds partially in the excited state and continues in the ground state, after which the chromophore follows one of two possible pathways: It breaks its hydrogen bond and forms the I_0 intermediate, or the hydrogen bond remains intact and the chromophore re-isomerizes to the ground state. The isomerization process and the breaking of the hydrogen bond are therefore independent processes. We propose that the strength of the hydrogen bond with Cys69 determines whether or not the chromophore will enter the photocycle. Upon light absorption, the hydrogen bond is weak enough to be broken in only ~30% of the protein molecules. In the remaining fraction, the chromophore relaxes to its stable ground state in ~6 ps. The strength of the hydrogen bond could determine the isomerization mechanism (using the multiple-bond flip to form I_0 or the hula-twist mechanism to form the ground state intermediate), but this is not unambiguously shown by these results. Changing the H-bond donating residue Glu46 near the phenolic oxygen with the weaker H-bond donating Gln-residue hardly alters the kinetics and product yields, even though during the I_0 to I_1 transition the hydrogen bond

of the chromophore with Gln46 is broken, while it remains intact with Glu46. The hydrogen bond with Glu46 therefore does not seem to play a crucial role in the initial part of the photocycle.

Acknowledgements

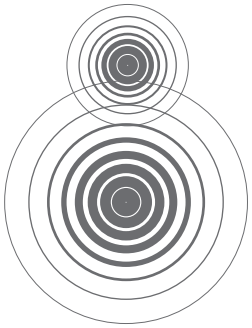
This research was supported by The Netherlands Organization for Scientific Research via the Dutch Foundation for Earth and Life Sciences (Investment Grant 812.08.001 and MtC-grant 805.47.123 to K.J.H., long-term fellowship 834.01.002 to M.L.G). L.J.G.W.v.W. is grateful to J. Key for carefully reading the manuscript.

References

1. Hellingwerf, K. J., Hendriks, J. & Gensch, T. (2003) *J. Phys. Chem. A* 107, 1082-1094.
2. Cusanovich, M. A. & Meyer, T. E. (2003) *Biochemistry* 42, 4759-4770.
3. Larsen, D. S. & van Grondelle, R. (2005) *Chem. Phys. Chem.* 6, 828-837.
4. Chosrowjan, H., Mataga, N., Nakashima, N., Yasushi, I. & Tokunaga, F. (1997) *Chem. Phys. Lett.* 270, 267-272.
5. Mataga, N., Chosrowjan, H., Shibata, Y., Imamoto, Y. & Tokunaga, F. (2000) *J. Phys. Chem. B* 104, 5191-5199.
6. Brudler, R., Rammelsberg, R., Woo, T. T., Getzoff, E. D. & Gerwert, K. (2001) *Nat. Struct. Biol.* 8, 265-70.
7. Xie, A., Hoff, W. D., Kroon, A. R. & Hellingwerf, K. J. (1996) *Biochemistry* 35, 14671-14678.
8. Imamoto, Y., Shirahige, Y., Tokunaga, F., Kinoshita, T., Yoshihara, K. & Kataoka, M. (2001) *Biochemistry* 40, 8997-9004.
9. Genick, U. K., Soltis, S. M., Kuhn, P., Canestrelli, I. L. & Getzoff, E. D. (1998) *Nature* 392, 206-9.
10. Perman, B., Srajer, V., Ren, Z., Teng, T., Pradervand, C., Ursby, T., Bourgeois, D., Schotte, F., Wulff, M., Kort, R., Hellingwerf, K. & Moffat, K. (1998) *Science* 279, 1946-50.
11. Ihee, H., Rajagopal, S., Srajer, V., Pahl, R., Anderson, S., Schmidt, M., Schotte, F., Anfinrud, P. A., Wulff, M. & Moffat, K. (2005) *Proc. Natl. Acad. Sci. USA* 102, 7145-7150.
12. Borgstahl, G. E., Williams, D. R. & Getzoff, E. D. (1995) *Biochemistry* 34, 6278-87.
13. Dux, P., Rubinstenn, G., Vuister, G. W., Boelens, R., Mulder, F. A., Hard, K., Hoff, W. D., Kroon, A. R., Crielaard, W., Hellingwerf, K. J. & Kaptein, R. (1998) *Biochemistry* 37, 12689-99.
14. Premvardhan, L. L., van der Horst, M. A., Hellingwerf, K. J. & van Grondelle, R. (2003) *Biophys. J.* 84, 3226-3239.
15. Larsen, D. S., van Stokkum, I. H. M., Vengris, M., van der Horst, M. A., de Weerd, F. L., Hellingwerf, K. J. & van Grondelle, R. (2004) *Biophys. J.* 87, 1858-1872.
16. Ujj, L., Devanathan, S., Meyer, T. E., Cusanovich, M. A., Tollin, G. & Atkinson, G. H. (1998) *Biophys. J.* 75, 406-12.
17. Baltuška, A., van Stokkum, I. H. M., Kroon, A., Monshouwer, R., Hellingwerf, K. J. & van Grondelle, R. (1997) *Chem. Phys. Lett.* 270, 263-266.
18. Devanathan, S., Pacheco, A., Ujj, L., Cusanovich, M., Tollin, G., Lin, S. & Woodbury, N. (1999) *Biophys. J.* 77, 1017-23.
19. Imamoto, Y., Kataoka, M., Tokunaga, F., Asahi, T. & Masuhara, H. (2001) *Biochemistry* 40, 6047-6052.
20. Gensch, T., Gradinaru, C. C., van Stokkum, I. H. M., Hendricks, J., Hellingwerf, K. J. & van Grondelle, R. (2002) *Chem. Phys.*

Lett. 356, 347-356.
21. Groot, M.-L., van Wilderen, L. J. G. W., Larsen, D. S., van der Horst, M. A., van Stokkum, I. H. M., Hellingwerf, K. J. & van Grondelle, R. (2003) *Biochemistry* 42, 10054-10059.
22. Heyne, K., Mohammed, O. F., Usman, A., Dreyer, J., Nibbering, E. T. J. & Cusanovich, M. A. (2005) *J. Am. Chem. Soc.* 127, 18100-18106.
23. Imamoto, Y., Kataoka, M. & Tokunaga, F. (1996) *Biochemistry* 35, 14047-53.
24. Taylor, R. & Kennard, O. (1984) *Acc. Chem. Res.* 17, 320-326.
25. Chosrowjan, H., Mataga, N., Shibata, Y., Imamoto, Y. & Tokunaga, F. (1998) *J. Phys. Chem. B* 102, 7695-7698.
26. Sugishima, M., Tanimoto, N., Soda, K., Hamada, N., Tokunaga, F. & Fukuyama, K. (2004) *Acta Crystallogr. Sect. D* 60, 2305-2309.
27. Anderson, S., Srajer, V. & Moffat, K. (2004) *Photochem. Photobiol.* 80, 7-14.
28. Genick, U. K., Devanathan, S., Meyer, T. E., Canestrelli, I. L., Williams, E., Cusanovich, M. A., Tollin, G. & Getzoff, E. D. (1997) *Biochemistry* 36, 8-14.
29. van Brederode, M. E., Gensch, T., Hoff, W. D., Hellingwerf, K. J. & Braslavsky, S. E. (1995) *Biophys. J.* 68, 1101-9.
30. van Stokkum, I. H. M., Larsen, D. S. & van Grondelle, R. (2004) *Biochim. Biophys. Acta* 1657, 82-104.
31. Zhou, Y., Ujj, L., Meyer, T. E., Cussanovich, M. A. & Atkinson, G. H. (2001) *J. Phys. Chem. A* 105, 5719-5726.
32. Kim, M., Mathies, R. A., Hoff, W. D. & Hellingwerf, K. J. (1995) *Biochemistry* 34, 12669-72.
33. Harada, I. & Takeuchi, H. (1996) in *Spectroscopy of Biological Systems*, eds. Clark, R. J. H. & Hester, R. E. (John Wiley & Sons, Chichester, UK), Vol. 13.
34. van Thor, J. J., Pierik, A. J., Nugteren-Roodzant, I., Xie, A. H. & Hellingwerf, K. J. (1998) *Biochemistry* 37, 16915-16921.
35. Xie, A., Kelemen, L., Hendriks, J., White, B. J., Hellingwerf, K. J. & Hoff, W. D. (2001) *Biochemistry* 40, 1510-7.
36. Unno, M., Kumauchi, M., Sasaki, J., Tokunaga, F. & Yamauchi, S. (2002) *Biochemistry* 41, 5668-5674.
37. Groenhof, G., Lensink, M. F., Berendsen, H. J. C., Snijders, J. G. & Mark, A. E. (2002) *Proteins Struct. Funct. Genet.* 48, 202-211.
38. Ren, Z., Perman, B., Srajer, V., Teng, T.-V., Pradervand, C., Bourgeois, D., Schotte, F., Ursby, T., Kort, R., Wulff, M. & Moffat, K. (2001) *Biochemistry* 40, 13788-13801.
39. Yeremenko, S., van Stokkum, I. H. M, Moffat, K. & Hellingwerf, K. J. (2006) *Biophys. J*. 90, 4224-4235.
40. van der Horst, M. A., van Stokkum, I. H. M., Dencher, N. A. & Hellingwerf, K. J. (2005) *Biochemistry* 44, 9160-9167.

41. Kort, R., Hellingwerf, K. J. & Ravelli, R. B. G. (2004) *J. Biol. Chem.* 279, 26417-26424.
42. Takeshita, K., Imamoto, Y., Kataoka, M., Mihara, K., Tokunaga, F. & Terazima, M. (2002) *Biophys. J.* 83, 1567-1577.
43. Groenhof, G., Bouxin-Cademartory, M., Hess, B., de Visser, S. P., Berendsen, H. J., Olivucci, M., Mark, A. E. & Robb, M. A. (2004) *J. Am. Chem. Soc.* 126, 4228-4233.
44. Imamoto, Y., Kataoka, M. & H., L. R. S. (2002) *Photochem. Photobiol.* 76, 584-589.
45. Hendriks, J., Gensch, T., Hviid, L., van der Horst, M. A., Hellingwerf, K. J. & van Thor, J. J. (2002) *Biophys. J.* 82, 1632-1643.
46. Marcano, A. O., Rodriguez, L. & Alvarado, Y. (2003) *J. Opt. A* 5, S256-S261.
47. Kurian, A., Unnikrishnan, K. P., George, D. S., Gopinath, P., Nampoori, V. P. N. & Vallabhan, C. P. G. (2003) *Spectrochim. Act. A* 59, 487-491.
48. Hoff, W. D., van Stokkum, I. H. M., van Ramesdonk, H. J., van Brederode, M. E., Brouwer, A. M., Fitch, J. C., Meyer, T. E., van Grondelle, R. & Hellingwerf, K. J. (1994) *Biophys. J.* 67, 1691-1705.



Dynamics of Carbon Monoxide Photodissociation in Bradyrhizobium Japonicum FixL Probed by Picosecond Mid-infrared Spectroscopy

05

L.J.G.W. van Wilderen*, J. Key*, I.H.M. Van Stokkum, R. van Grondelle, M.L. Groot
***Authors contributed equally to this work**
Manuscript in preparation

Abstract

The FixL proteins are heme-based bacterial oxygen sensors, distinct from globins in structure and ligand binding properties. In order to better understand the dynamics of ligand binding within the PAS domain fold, we have carried out picosecond visible pump - mid-infrared probe spectroscopy. We used the diatom CO as a ligand to the heme, which upon photoexcitation with a visible laser pulse is released. The infrared-active stretch frequency of the CO molecule is very sensitive to interactions with the surrounding protein, and acts therefore as a probe for the chromophore-binding pocket dynamics. The infrared absorption-difference spectra indicate that the escape of photolyzed CO to solvent is preceded by transient docking within the protein in a manner similar to globins. On a picosecond timescale protein rearrangements lead to a decrease of the rotational constrainedness of the CO molecule. On a nanosecond timescale further protein relaxations occur that either cause the CO molecule to migrate to a new site, or to change its rotational orientation, and create a more heterogeneous environment. The loosening of the protein around CO in FixL is a marked difference with the tightening of the protein that occurs around CO in myo- and hemoglobin. Most likely these events are the onset of the signal transduction process in FixL.

Introduction

The PAS domain is a ubiquitous ~100 amino acid signaling module found in proteins from all kingdoms of life. These domains are often coupled to an enzymatic output domain whose activity is regulated by the PAS domain in response to an external stimulus. Proteins containing PAS domains exhibit immense diversity; they include photoreceptors (1, 2), redox sensors (3, 4), and PAS domains which bind small molecules (5). PAS domains can contain a variety of prosthetic groups, though most do not. PAS domain proteins which contain a heme prosthetic group have emerged as an important subgroup of the PAS super family. These are distinct, in that they contain a covalently-bound protoporphyrin IX (heme) prosthetic group, and typically exhibit catalytic control of their enzymatic domains in response to dissolved diatomic gases. Heme-PAS proteins have been identified from bacteria (6) to mammals (7) (see (8) for a review). The first and best characterized of these proteins are the FixL proteins of nitrogen-fixing *Rhizobia*.

Bradyrhizobium japonicum FixL (hereafter referred to as *Bj*FixL) is a dimeric oxygen sensor protein responsible for regulation of the transcription factor FixJ in response to oxygen. FixJ, in turn, regulates expression of proteins involved in nitrogen fixation and oxidative stress. *Bj*FixL is composed of two domains; an N-terminal heme-PAS sensory domain and a C-terminal kinase domain. Ligand binding at the heme of the N-terminal domain results in inhibition of autophosphorylation in the C-terminal domain. Numerous crystallographic studies (9-12) have focused on the isolated heme-PAS domain of this protein, referred to as *Bj*FixLH (Figure 1). These studies have revealed that, upon binding of oxygen to deoxy *Bj*FixLH, a conformational change occurs in the FG loop and in the H and I strands of the β sheet of the protein (11). The side chain of Arg220 moves into the heme pocket where it serves as a distal coordinating residue to the oxygen ligand. Carbon monoxide, however, does not elicit this change in Arg220, yet structural changes are evident in the β strands of the protein, a region implicated in signal transduction in other PAS domains of diverse function (12). The mechanism by which ligand binding regulates the C-terminal kinase is unclear.

The diatomic heme ligand carbon monoxide has been used as a molecular probe for protein dynamics in hemoglobin (Hb) and myoglobin (Mb) for many years (13-16). The heme Fe-CO bond is broken upon electronic excitation of the heme with high quantum efficiency, rapidly generating a non-equilibrium state within the protein whose relaxation can be monitored experimentally. In globins, photolysis of the CO-Fe bond is followed by release of the CO molecule to the solvent and concomitant relaxation of the surrounding protein (17). De-ligated CO follows a defined trajectory through the protein to the solvent via several internal protein binding sites (18, 19). These sites are hydrophobic in nature and

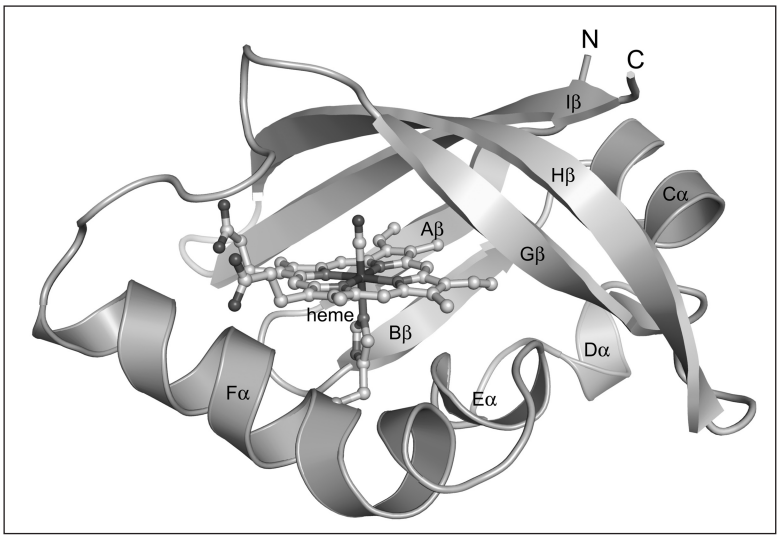


Figure 1 **Schematic representation of *Bj*-FixLH, drawn as a ribbon model, with the central heme group drawn in balls-and-sticks. The CO-ligand is bound to the heme, pointing upward from the plane of the heme, with the oxygen drawn in black.**

have been referred to as xenon sites for their propensity to bind a xenon atom (20). However, little is known about the dynamics of CO photolysis within the widely-distributed heme-PAS domain which differs substantially from globins in structure and ligand binding properties (6). Earlier studies using UV-visible femtosecond spectroscopy noted no rebinding of CO within 4 ns, suggesting the photodissociated CO proceeds efficiently to solvent in wild-type protein (21) and geminate CO recombination is only observed in the presence of a distal ligating residue (22).

We present picosecond visible pump mid–infrared probe spectroscopy on the CO-*Bj*FixLH complex in order to reveal the dynamics associated with the escape of CO from the protein into the solvent. Changes in the oscillator strength and stretch frequency of CO, due to changes in the interaction between free CO and its protein/solvent environment, have been monitored spectroscopically from 150 fs to 3 ns after photolysis of the CO.

Materials and methods

Protein preparation

*Bj*FixLH was purified as described in (23). The sample was prepared in H_2O at a concentration of ~10 mM *Bj*FixLH in 200 mM NaCl, 50 mM Tris pH 8. Dithionite was used to reduce the heme protein solution, and was immediately followed by flushing with CO gas to saturate the unligated heme with the diatom. The sample consisted of a protein solution pressed between two CaF_2 windows, separated by a Teflon spacer with a thickness of 100 μm. This resulted in an absorption at 542 nm of about 0.7 OD and an OD of about 0.85 at 1966 cm^{-1}. UV-VIS and IR spectra of the sample collected before and after each measurement were used to monitor the binding of CO to the heme (see Figure 2 for a typical

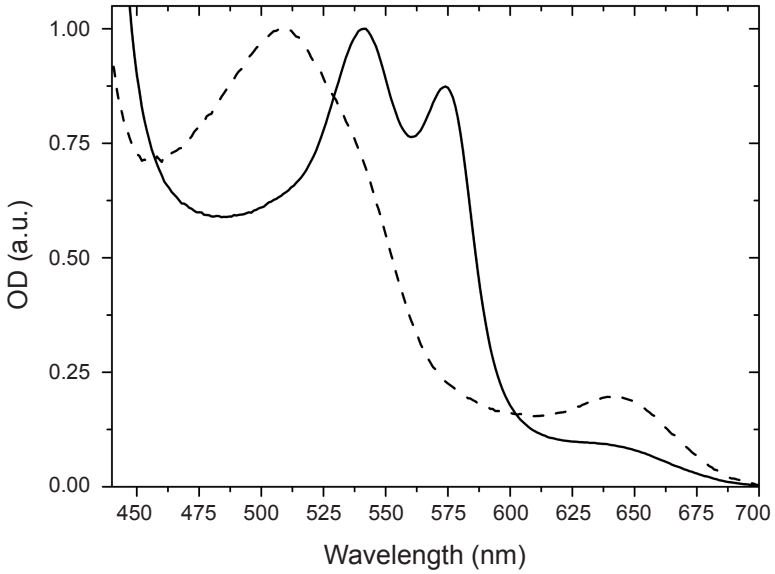


Figure 2 **Absorption spectrum of *Bj*FixLH with CO bound to the heme (continuous) and with the heme in the oxidized state (dashed), normalized on the peak in the 500-550 nm region.**

absorption spectrum of FixL with the heme in the CO-ligated and unligated state). No apparent changes were observed in either spectrum due to the photolysis experiment.

Experimental setup

Visible absorption spectra were measured on a UV–VIS spectrometer (Lambda 40, Perkin–Elmer, Wellesley, MA, USA). Infrared absorption spectra were measured with a FTIR spectrometer (Avatar 370, Thermo Nicolet Waltham, MA, USA). The experimental setup (see (24) and chapter 1 for a more detailed description) for the infrared difference absorption measurements consisted of an integrated Ti:sapphire oscillator/ regenerative amplifier, operating at 1 kHz, and producing 0.8 mJ pulses of 80 fs (Hurricane, SpectraPhysics, Mountain View, CA, USA). The output of this laser was used to pump a commercial optical parametric generator and amplifier with difference frequency generation (TOPAS, Light Conversion, Vilnius, Lithuania), which resulted in a tuneable output (2.5–10 μm) with a spectral width of ~200 cm^{-1}. A home–built HgCdT camera system placed behind a spectrograph was read out every shot at a repetition rate of 1 kHz and a sampling resolution of ~4–5 cm^{-1}.

Two excitation wavelengths were used to initiate the photolysis, *i.e.* 402 nm and 542 nm. The first pump–wavelength was generated by frequency doubling a part of the Hurricane output in a BBO crystal, producing pulses at 402 nm with a power of 250 nJ. The 800–nm light that was not converted by the crystal is being filtered out. Alternatively, for 542–nm excitation, the output of the Hurricane that was used to generate 402–nm light was now used to pump a home-built non-collinear OPA. The pulses had a duration of ~60 fs (uncompressed) and an energy of 140-800 nJ.

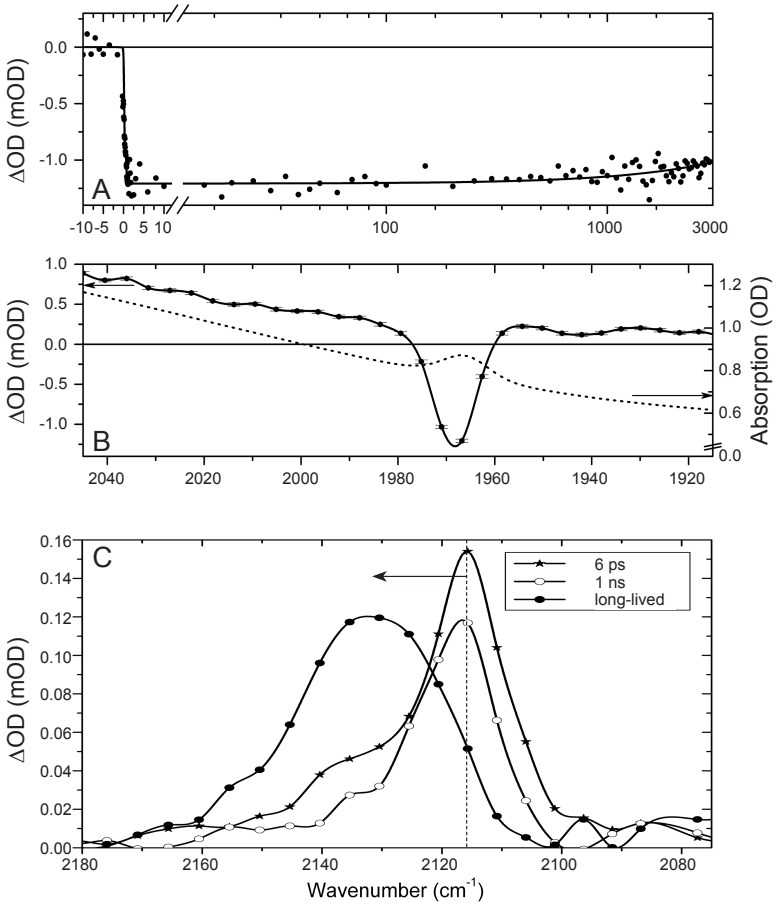


Figure 3 **Results of pump-probe measurements on *Bj*FixLH, excited with 542 nm. In panel A, the photoinduced maximum absorption change at 1966 cm^{-1} is shown as a function of time. The time axis is linear up to 10 ps and logarithmic thereafter. Panel B shows the results of the global analysis (EADS, with error bars) of the bound CO region (solid line), with the steady state FTIR spectrum of CO bound *Bj*FixLH for comparison (dashed). Panel C contains the background subtracted-EADS of the average of all individual EADS's (containing four data-sets with a total of 248 scans) in the spectral region where the free CO is observed. The dashed line marks the maximum difference signal on a nanosecond time scale. Only those components with a similar time constant were averaged.**

A moveable delay line (double pass) made it possible to increase the time–of–arrival of the pump–beam to 3 ns with respect to the probe–beam. The pump beam polarization was set with a Berek–rotator (New Focus 5540, San Jose, CA, USA) to magic angle with respect to the probe beam. A phase–locked chopper at 500 Hz ensured that every other shot the sample was excited and an absorbance difference spectrum could be calculated. To ensure a fresh spot for each laser shot, the sample was moved with a homebuilt Lissajous scanner. The setup was contained in a dry–air–purged box to reduce distortions of the infrared beam by water vapour absorption.

Data was collected over two spectral windows: from 1915–2045 cm^{-1} and from 2077–2180 cm^{-1}, with spectral resolutions of ~4 and ~5 cm^{-1}, respectively. To check reproducibility, every window was recorded during three different periods, and using three different batches of sample. A measurement session typically consisted of 30 – 60 scans (referred to as one *dataset*), and took about 1.2–2.5 hours. The data was recorded with homebuilt software, and analyzed with the global analysis method (25). A correction was

made for the presence of a pre–time–zero offset due to thermal lens effects (26, 27). For each excitation wavelength, the resulting evolution associated difference spectra (EADS) that had similar time constants were averaged. The average EADS resulting from 542-nm excitation are shown in Figure 3 and drawn with a spline function to represent more realistic spectral bands. Figures 3A and B were the result of 80 scans, while Figure 3C is the result of 248 scans, resulting in a signal-to-noise level of about 1300 μOD and 300 μOD, respectively. Because of the presence of a small perturbed free–induction decay before $t = 0$, no information faster than the instrument response (~180 fs) was extracted.

Results

Absorption changes in spectral windows containing bands associated with bound CO (1915–2045 cm^{-1}) and photolyzed CO (2077–2180 cm^{-1}) were recorded, each pumped with 542 and 402–nm light to initiate the release of CO. Both excitation wavelengths reveal similar results, but because of the higher quality of the 542-nm excitation data only these data are shown. The photo-induced absorption changes relative to the ground state are measured as a function of time; positive peaks are due to product formation, negative peaks represent the ground state. Spectral changes within each window are made visible with the aid of global analysis, in which all wavelengths are fitted simultaneously (25), resulting in EADS.

Ligated CO

The peak of the recorded absorption change within the 1915–2045 cm^{-1} spectral window (Figure 3B) is at ~1966 cm^{-1} , both upon 542-nm and 402-nm excitation, and coincides with a band in the steady–state FTIR absorption spectrum (dashed spectrum in Figure 3B) which is only present when CO is bound to the heme. This feature is therefore due to the bleached absorption of bound CO. The frequency is similar to that found for the heme proteins Hb and Mb, which show a bleach upon photodissociation at 1952 (28) and 1947 cm^{-1} (29), respectively. The kinetics associated with the band at ~1966 cm^{-1} - constant throughout the measured 3–ns time range - indicate that CO is released upon photolysis and does not recombine on the timescale of the experiment (3 ns)*, in agreement with earlier visible pump–probe experiments (21).

Dissociated CO

The CO molecule shows an upshift in frequency to the 2077–2180 cm^{-1} region of the spectrum, upon breaking of the CO-Fe bond. Global analysis of this spectral window (Figure 3C)

* A small recovery of the bleach can be seen at long delay times (from 2.8 to 3 ns for the 542-nm excitation), but this is an artefact of the delay line alignment.

reveals a positive band around 2116 cm^{-1} with a FWHM of ~16 cm^{-1}, appearing simultaneously with the bleach of bound CO at ~1966 cm^{-1} (Figure 3A). In addition to the 2116 cm^{-1} band, a distinct shoulder can be observed on the high frequency side (at 2135-2140 cm^{-1})†. On a picosecond- and nanosecond time scale dynamics in the product bands are observed, concurrent with relatively large changes in background absorption (see Figure 5 of the Supplementary Information, from now on referred to as SI). The latter are attributed to sample heating effects (see below), and to emphasize the spectral changes related to CO, a linear baseline (see Figure 5 of the SI) has been subtracted from the averaged EADS (30-32). This procedure reveals that, with a time constant of ~6 ps, the photodissociated CO band decreases in intensity, and shows a small blue shift. A further blue shift occurs with a time constant of ~1 ns, resulting in a broadened single band‡ peaking at 2135 cm^{-1}, with a FWHM of ~28 cm^{-1}. On both sides of the main band two small shoulders appear, but these are on the limit of the signal–to–noise.

Discussion

The recorded infrared difference spectra of bound and free CO will be discussed, and compared with earlier reported results on FixL, Hb and Mb. A virtually inevitable effect, *i.e.* heating of the protein upon laser excitation, will be discussed first.

Heating effects

Upon photoexcitation, a large amount of energy is deposited in the ligand-bound protein, leading to an increase of the temperature of the heme. The temperature can, depending on the choice of excitation wavelength, rise with more than 250 K (33, 34). The excess energy can be dissipated in a number of ways. The heme can release this energy by energy transfer to the ligand (instantaneously), or vibrational cooling of the heme (a 6.2-ps process in Mb (34), ultimately leading to heating of the solvent. The heme (in globins) can also heat the solvent more directly by specific motions of its proprionic side chains that are exposed to the solvent (a ~9 ps process, built up from a combination of heme cooling in ~6 ps and water heating in ~3-4 ps (35)). Dissolved in water, cooling of the hot heme will thus lead to an increase of the vibrational frequency of hydrogen bonds (water at room temperature has a wide absorption peak in the studied mid-infrared region). Heating effects assigned to water have been reported for Hb and Mb (14,

† The 402-nm excitation reveals, apart from the same high-frequency shoulder, also a low-frequency shoulder at 2087 cm^{-1}.

‡ When 402-nm excitation is used, the central band at 2116 cm^{-1} upshifts to a double peaked 2125-2135 cm^{-1} band (with the 2125 cm^{-1} band more intense). The additional low frequency shoulder (at 2087 cm^{-1}) upshifts to 2106 cm^{-1}.

35). The end result, in terms of the picosecond-midIR experiment, is a dynamic change in background absorption. A recent extensive water heating study illustrates that the absorption in the 1900–2400 cm^{-1} spectral region (36) can change upon transfer of laser pulse energy (532 nm, 2 mJ/cm^2) directly to the water solvent, resulting in a temperature increase of 0.2ºC of the water. It is known that the spectral shift of H_2O is linear with the temperature in the range of 300 K (37). The resulting change in optical density of water (+0.1 below 2050 cm^{-1} to –0.3 mOD above 2050 cm^{-1}) is identical to changes measured for bacteriorhodopsin in water (36). Likewise, our results reveal the start of water heating dynamics. A positive offset is evident in the 1915–2045 cm^{-1} CO bleach region (Figure 3A), a negative one in the 2066–2166 cm^{-1} region (see Figure 5 of the SI). Similar studies with a dye (malachite green) and heme protein (Hb and Mb) solutions showed that the heating appeared with two time constants. A fast 8-ps component was suggested to be caused by low frequency motions of the protein that facilitate the energy transfer to the solvent, or direct heating of the solvent through the proprionic side chains of the heme, or a combination of the two; A slow 20-ps phase was attributed to an energy transfer process from the heme to water via a diffusion process (35). Also, no changes were observed from 100 ps to 2 ns, indicating that the average water temperature is constant during this period (35). The 6–ps lifetime we observe is therefore most likely related to cooling of the heme and the heating of water.

Bound CO

Upon photodissociation with both 402-nm and 542-nm excitation, a single instantaneous bleach is observed. This bleach is located at ~1966 cm^{-1}, has a FWHM of ~10 cm^{-1}, and does not show any spectral evolution, indicating that the CO does not rebind to the heme on a 3-ns time scale (Figure 3A). This observation is similar to a previous visible flash photolysis study of *Bj*FixLH, which reported no rebinding for more than 4 ns (38). For Hb and Mb, the CO-photodissociation yield is unity (30), and the absence of rebinding on the pico- or nanosecond time scale suggests that this is the case too for *BjFixLH*. In Mb, multiple bound CO bands have been observed, which, since their anisotropy is the same, arise from conformational substates with different electrostatic fields in the vicinity of the active binding site, rather than from different CO orientations (39). Bound CO conformers in Mb have FWHM bandwidths ranging from 8 to 18 cm^{-1} (40, 41). Our wavelength resolution is about 4 cm^{-1} in this spectral region, but we do not see evidence of other peaks at room temperature. Thus, we conclude that in CO-*Bj*FixLH the active site exists in only one conformation.

Carbon monoxide recombination in globins is characterised by a unimolecular geminate phase in which a photolyzed CO

molecule rebinds to the heme from which it was photodissociated, and a slower bimolecular phase, where CO escapes to and rebinds from the solvent (42). The existence of a transient docking site in *Bj*FixLH would suggest a geminate phase, and a unimolecular equilibrium between the transiently docked CO and the newly vacated heme should be evident. However, our results (Figure 3A), as well as others (21), reveal no geminate phase. In *Bj*FixLH, geminate recombination of CO has only been observed in mutants constructed with a distal Histidine residue, analogous to the distal residue in most globins. Analogously, Mb mutants lacking a distal coordinating residue have substantially diminished geminate recombination (43, 44), yet CO docking is evident (45, 46). This suggests that the character of this distal residue is a greater influence on geminate recombination than the existence of docking sites.

Photodissociated CO

The induced absorption signal from the photolyzed CO upshifts to higher frequencies (Figure 3C) and gains a factor 1.1 in oscillator strength on a 1-ns time scale. The ratio of integrated intensities of bound and photolyzed CO in Mb at 100 ps is reported to vary from 21:1 to 34:1, depending on temperature (47, 48). For the spectrum with the 6–ps lifetime, we find ratios of ~19:1 and ~15:1 for 542-nm and 402–nm excitation, respectively. The datasets recorded with 542 and 402–nm excitation each show qualitatively the same behaviour; an initial spectrum with one main peak at 2116 cm^{-1} and a shoulder at ~2135-2140 cm^{-1}, both decreasing in intensity with a lifetime of 6 ps, and a blue shift on the nanosecond time scale to 2135 cm^{-1}.

At first instance one might wonder whether the two bands in the t = 0 spectra reflect CO molecules in vibrationally excited states. The bands have an energy difference of 19-24 cm^{-1}, similar to the energy difference between the ω_{10} and ω_{21} bands of CO of 25 cm^{-1} for Hb and Mb (49-51). However, relaxation to the vibrational lowest energy state, *i.e.* at 2135 cm^{-1}, would then only occur on the 1-ns time scale, whereas vibrational population relaxation has been reported to take ~25 ps for the I_1 to I_0 transition (51, 52). In addition, the high intensity of the low-frequency band would suggest an extremely high temperature of the CO molecule. The two bands are therefore more likely to represent the B_1 and B_2 states that in Mb and Hb reflect the CO molecule in different environments. Indeed, the frequencies of the observed bands in the photolyzed spectrum (*i.e.* at 2135 and 2116 cm^{-1}) are very similar to those observed for the B_1 and B_2 states in Mb (reported at 2132 and 2122 cm^{-1}, respectively (41, 48)). The two B bands in Mb have been proposed to correspond to the same protein pocket occupation of the CO molecule, but in a 'head first' or 'tail first' orientation, each configuration interacting with its environment

in a different way, hence the different frequencies (53, 54). This suggests that also in *Bj*FixLH there are two different conformations of CO, with a preferential population of the 2116 cm^{-1} band. The relative constant population of the 2116 and 2135 cm^{-1} bands at t = 0 and after 6 ps, suggests that they do not interconvert on the picosecond time scale. The width of the B_1 and B_2 bands (12-16 cm^{-1} when fitted with two Gaussians, similar to the B_1 and B_2 bands in Mb which have widths of ~11 and ~20 cm^{-1}, respectively (48)) contains information about the interaction between the CO molecule and its surroundings, and its narrowness indicates that the CO is still in a constrained environment. The total difference absorbance signal looses intensity and shows a minor blue shift on a 6-ps time scale. The loss of peak intensity does not reflect population decay, as we see no corresponding recovery of the bleached bound-CO band. A loss of signal may reflect a population of CO molecules that escapes the protein into the solvent already on the 6-ps time scale. In the solvent, the line width of CO is much broader (55) and, assuming the integrated oscillator strength is more or less constant, hence would be expected to lead to a loss of signal. The fact that signal is gained again on the 1-ns time scale however rules out this possibility. In Mb, an *increase* of signal on the 1.6-ps time scale was observed (53). Lim *et al* (48) suggested that the integrated absorbance of CO is partitioned between the narrow B-states and a broad (~90 cm^{-1}) unresolved feature. The partitioning is determined by the orientational constraints imposed on CO by the protein, and the signal increase was consequently interpreted to be caused by a relaxation of the protein around CO (thereby limiting its orientational freedom). In *Bj*FixLH, the *loss* of signal indicates rather a loosening of the protein pocket around CO on the 6-ps time scale, giving the CO more rotational freedom. Unfortunately, due to the anticipated width of the less-constrained CO band (~90 cm^{-1}) and the interference with the heated-water background, the full line shape of the CO molecule is not resolved, and we can only observe the narrow part of the line shape.

The time constant of this loosening process is the same as we observe for the heating of water, and similar to the vibrational cooling time reported for Mb (6.2 ps in H_2O (34), 7.5 ps in D_2O (35)) and Hb (8.5 ps in D_2O (35)). This may suggest that loosening of the protein pocket is induced by the cooling of the heme, or that the environment of the CO molecule is still dominated by the heme because of its close proximity to the heme. Integration of each spectrum over the measured spectral window, yields 3.8, 2.5, and 4.1 mOD·cm^{-1} for the 6–ps, 1–ns, and long–lived spectra, respectively. The FWHM changes, however, only with the long–lived component (from 16 to 28 cm^{-1}). On a nanosecond time scale, spectral changes occur that may be interpreted in two ways; both the B_1 and B_2 state seem to upshift with ~20 cm^{-1}, suggesting either a further relaxation of the protein, or the migration of the

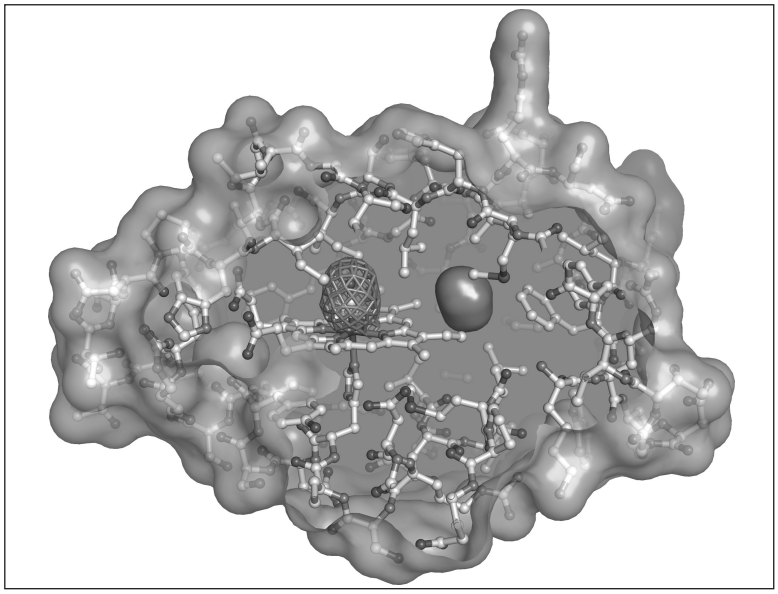


Figure 4 **Schematic drawing of a slice through *Bj*FixLH, revealing the heme-binding pocket inside the protein. The protein surface is transparent. The CO molecule bound to the heme is drawn with a mesh structure (revealing the CO molecule in balls-and-sticks inside). *Bj*FixLH, with the CO molecule photodissociated from the heme, reveals a cavity near the heme (symbolized with the dark grey solid shape to the right of the heme).**

CO molecule to a different docking site; Or the small 2155 cm^{-1} shoulder is not a real feature, and a nanosecond interconversion process from the initially more preferred orientation (2116 cm^{-1}) of the CO molecule to the less preferred orientation (2135 cm^{-1}) occurs, again induced by protein rearrangements. Note also that the main peak in the long-lived spectrum (peaking at 2135 cm^{-1}) is very wide, suggesting more heterogeneity in the CO frequency.

A search for cavities within the *Bj*FixLH protein structure akin to those found in globins reveals only one small cavity within the core of the PAS domain bordered by the periphery of the porphyrin ring and the side chains of Phe176 and Met134 (see Figure 4). This strengthens our hypothesis that the two bands in the photolyzed CO spectrum reflect the CO molecule in two different orientations within one protein cavity. The differences we observe between Mb and *Bj*FixLH, *i.e.* a tightening of the protein pocket versus a relaxation of the protein pocket, is likely a reflection of the signal transduction function of *Bj*FixLH.

Conclusions

Probing the vibrational frequency of the bound and free CO molecule upon photoexcitation provides a large amount of information on protein properties. In *Bj*FixLH we find that the escape of photolyzed CO to solvent is preceded by transient docking within the protein in a manner similar to globins. This docking appears to involve two orientations of the CO. On a 6-ps time scale we observe a loss of signal in the spectral region of the photolyzed CO, which we ascribe to a loosening of the environment around the CO molecule. Further protein relaxation, as probed by

absorption changes of the photolyzed CO molecule, occurs on a 1-ns time scale. These protein structural changes constitute most likely the onset of the signal transduction in *Bj*FixLH.

Acknowledgement
The authors thank R. Schmitz for supplying the CO.

References

1. Borgstahl, G. E., Williams, D. R. & Getzoff, E. D. (1995) *Biochemistry* 34, 6278-87.
2. Crosson, S. & Moffat, K. (2001) *Proc Natl Acad Sci U S A* 98, 2995-3000.
3. Martinez-Argudo, I., Little, R., Shearer, N., Johnson, P. & Dixon, R. (2004) *J Bacteriol* 186, 601-10.
4. Gauden, M., van Stokkum, I. H., Key, J. M., Luehrs, D., van Grondelle, R., Hegemann, P. & Kennis, J. T. (2006) *Proc Natl Acad Sci U S A* 103, 10895-900.
5. Da Silva Xavier, G., Rutter, J. & Rutter, G. A. (2004) *Proc Natl Acad Sci U S A*.
6. Delgado-Nixon, V. M., Gonzalez, G. & Gilles-Gonzalez, M. A. (2000) *Biochemistry* 39, 2685-91.
7. Dioum, E. M., Rutter, J., Tuckerman, J. R., Gonzalez, G., Gilles-Gonzalez, M. A. & McKnight, S. L. (2002) *Science* 298, 2385-7.
8. Gilles-Gonzalez, M. A. & Gonzalez, G. (2005) *J Inorg Biochem* 99, 1-22.
9. Gong, W., Hao, B., Mansy, S. S., Gonzalez, G., Gilles-Gonzalez, M. A. & Chan, M. K. (1998) *Proc Natl Acad Sci U S A* 95, 15177-82.
10. Gong, W., Hao, B. & Chan, M. K. (2000) *Biochemistry* 39, 3955-62.
11. Hao, B., Isaza, C., Arndt, J., Soltis, M. & Chan, M. K. (2002) *Biochemistry* 41, 12952-8.
12. Key, J. & Moffat, K. (2005) *Biochemistry* 44, 4627-35.
13. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. (1975) *Biochemistry* 14, 5355-73.
14. Anfinrud, P. A., Han, C. & Hochstrasser, R. M. (1989) *Proc Natl Acad Sci U S A* 86, 8387-91.
15. Lim, M., Jackson, T. A. & Anfinrud, P. A. (1993) *Proc Natl Acad Sci U S A* 90, 5801-4.
16. Scott, E. E. & Gibson, Q. H. (1997) *Biochemistry* 36, 11909-17.
17. Brunori, M., Bourgeois, D. & Vallone, B. (2004) *J Struct Biol* 147, 223-34.
18. Schotte, F., Soman, J., Olson, J. S., Wulff, M. & Anfinrud, P. A. (2004) *J Struct Biol* 147, 235-46.
19. Hummer, G., Schotte, F. & Anfinrud, P. A. (2004) *Proc Natl Acad Sci U S A* 101, 15330-4.
20. Tilton, R. F., Jr., Kuntz, I. D., Jr. & Petsko, G. A. (1984) *Biochemistry* 23, 2849-57.
21. Liebl, U., Bouzhir-Sima, L., Negrerie, M., Martin, J. L. & Vos, M. H. (2002) *Proc Natl Acad Sci U S A* 99, 12771-6.
22. Balland, V., Bouzhir-Sima, L., Kiger, L., Marden, M. C., Vos, M. H., Liebl, U. & Mattioli, T. A. (2005) *J Biol Chem* 280, 15279-88.

23. Gilles-Gonzalez, M. A., Gonzalez, G., Perutz, M. F., Kiger, L., Marden, M. C. & Poyart, C. (1994) *Biochemistry* 33, 8067-73.
24. Groot, M. L., van Wilderen, L. J. G. W., Larsen, D. S., van der Horst, M. A., van Stokkum, I. H., Hellingwerf, K. J. & van Grondelle, R. (2003) *Biochemistry* 42, 10054-9.
25. van Stokkum, I. H., Larsen, D. S. & van Grondelle, R. (2004) *Biochim Biophys Acta* 1657, 82-104.
26. Marcano, A. O., Rodriguez, L. & Alvarado, Y. (2003) *Journal of Optics A: Pure and Applied Optics* 5, S256-S261.
27. Kurian, A., Unnikrishnan, K. P., George, D. S., Gopinath, P., Nampoori, V. P. N. & Vallabhan, C. P. G. (2003) *Spectrochimica Acta Part A* 59, 487-491.
28. Jedju, T. M., Rothberg, L. & Labrie, A. (1988) *Optics Letters* 13, 961-963.
29. Frauenfelder, H., McMahon, B. H., Austin, R. H., Chu, K. & Groves, J. T. (2001) *Proceedings of the National Academy of Sciences of the United States of America* 98, 2370-2374.
30. Chance, M. R., Courtney, S. H., Chavez, M. D., Ondrias, M. R. & Friedman, J. M. (1990) *Biochemistry* 29, 5537-5545.
31. Mourant, J. R., Braunstein, D. P., Chu, K., Frauenfelder, H., Nienhaus, G. U., Ormos, P. & Young, R. D. (1993) *Biophysical Journal* 65, 1496-1507.
32. Jackson, T. A., Lim, M. H. & Anfinrud, P. A. (1994) *Chemical Physics* 180, 131-140.
33. Henry, E. R., Eaton, W. A. & Hochstrasser, R. M. (1986) *Proceedings of the National Academy of Sciences of the United States of America* 83, 8982-8986.
34. Lim, M. H., Jackson, T. A. & Anfinrud, P. A. (1996) *Journal of Physical Chemistry* 100, 12043-12051.
35. Lian, T., Locke, B., Kholodenko, Y. & Hochstrasser, R. M. (1994) *Journal of Physical Chemistry* 98, 11648-11656.
36. Garczarek, F., Wang, J. P., El-Sayed, M. A. & Gerwert, K. (2004) *Biophysical Journal* 87, 2676-2682.
37. Draegert, D. A., Stone, N. W. B., Curnutte, B. & Williams, D. (1966) *Journal of the Optical Society of America* 56, 64.
38. Rodgers, K. R. & Lukat-Rodgers, G. S. (2005) *J Inorg Biochem* 99, 963-77.
39. Lim, M., Jackson, T. A. & Anfinrud, P. A. (2004) *Journal of the American Chemical Society* 126, 7946-7957.
40. Shimada, H. & Caughey, W. S. (1982) *Journal of Biological Chemistry* 257, 1893-1900.
41. Alben, J. O., Beece, D., Bowne, S. F., Doster, W., Eisenstein, L., Frauenfelder, H., Good, D., Mcdonald, J. D., Marden, M. C., Moh, P. P., Reinisch, L., Reynolds, A. H., Shyamsunder, E. & Yue, K. T. (1982) *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* 79, 3744-3748.

42. Carver, T. E., Olson, J. S., Smerdon, S. J., Krzywda, S., Wilkinson, A. J., Gibson, Q. H., Blackmore, R. S., Ropp, J. D. & Sligar, S. G. (1991) *Biochemistry* 30, 4697-4705.
43. Lambright, D. G., Balasubramanian, S., Decatur, S. M. & Boxer, S. G. (1994) *Biochemistry* 33, 5518-5525.
44. Brunori, M., Cutruzzola, F., Savino, C., Travaglini-Allocatelli, C., Vallone, B. & Gibson, Q. H. (1999) *Biophysical Journal* 76, 1259-1269.
45. Bourgeois, D., Vallone, B., Schotte, F., Arcovito, A., Miele, A. E., Sciara, G., Wulff, M., Anfinrud, P. & Brunori, M. (2003) *Proceedings of the National Academy of Sciences of the United States of America* 100, 8704-8709.
46. Bourgeois, D., Vallone, B., Arcovito, A., Sciara, G., Schotte, F., Anfinrud, P. A. & Brunori, M. (2006) *Proceedings of the National Academy of Sciences of the United States of America* 103, 4924-4929.
47. Polack, T., Ogilvie, J. P., Franzen, S., Vos, M. H., Joffre, M., Martin, J. L. & Alexandrou, A. (2004) *Physical Review Letters* 93, 01802.
48. Lim, M. H., Jackson, T. A. & Anfinrud, P. A. (1995) *Journal of Chemical Physics* 102, 4355-4366.
49. Rector, K. D., Kwok, A. S., Ferrante, C., Tokmakoff, A., Rella, C. W. & Fayer, M. D. (1997) *Journal of Chemical Physics* 106, 10027-10036.
50. Owrutsky, J. C., Li, M., Locke, B. & Hochstrasser, R. M. (1995) *Journal of Physical Chemistry* 99, 4842-4846.
51. Ventalon, C., Fraser, J. M., Vos, M. H., Alexandrou, A., Martin, J. L. & Joffre, M. (2004) *Proceedings of the National Academy of Sciences of the United States of America* 101, 13216-13220.
52. Lim, M. H., Hamm, P. & Hochstrasser, R. M. (1998) *Proceedings of the National Academy of Sciences of the United States of America* 95, 15315-15320.
53. Lim, M. H., Jackson, T. A. & Anfinrud, P. A. (1997) *Nature Structural Biology* 4, 209-214.
54. Meller, J. & Elber, R. (1998) *Biophysical Journal* 74, 789-802.
55. Ewing, G. E. (1962) *Journal of Chemical Physics* 37, 2250-2256.

Supplementary information

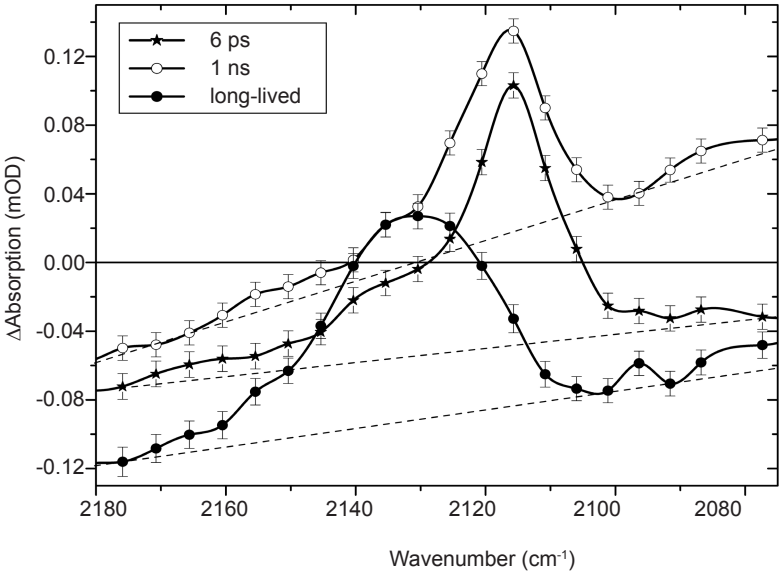
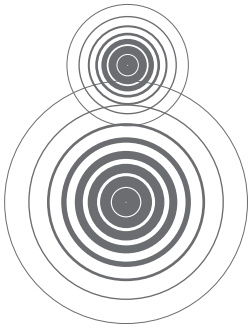


Figure 5 **EADS of the average of all individual EADS's (containing four datasets and 248 scans) in the spectral region where the free CO is observed, without background subtraction and with error bars. The data are identical to those presented in Figure 3C. Only those components with a similar time constant were averaged. Note the dynamics of the background of the spectra, which is modelled with a linear function for each spectrum (denoted with the dashed line under each spectrum it belongs to).**



Polarized Infrared Spectroscopy on Photoactive Yellow Protein

06

L.J.G.W. van Wilderen, I.H.M. Van Stokkum, M.A. van der Horst, K.J. Hellingwerf, R. van Grondelle, M.L. Groot

Manuscript in preparation

Abstract

Polarized infrared measurements on Photoactive Yellow Protein (PYP) are presented with the aim to resolve time–dependent changes in orientation of individual groups within the protein during the photoinduced isomerization reaction of the chromophore. We have measured pump–probe absorption difference spectra on PYP in D_2O with parallel, perpendicular and magic angle orientation of the polarization of the pump– and probe–laser pulses and calculated anisotropies for individual bands at time delays of 1 and 50 ps. The interpretation of the results obtained is hampered by the delocalized nature of many of PYP's vibrations, the signal–to–noise ratio of the data, and the presence of a mixture of states at the time delays used. We focus on the 1600–1780 cm^{-1} region which contains the C=O groups of the chromophore and of Glu46. The latter is H–bonded to the phenol moiety of the chromophore, while the chromophore's C=O group is H–bonded to the protein backbone. Our anisotropy data provide independent evidence for the hypothesis that, upon photoexcitation, the chromophore's C=O remains H–bonded in the excited state, with characteristics similar to those of a single bond. Upon formation of I_0, the first stable photocycle intermediate, the H–bond between Glu46 and the phenol moiety of the chromophore is confirmed to strengthen. The I_0 state is characterized by a broken H–bond between the chromo–phore's C=O and the backbone, and by a different orientation of the same C=O with respect to the ground state. Our results illustrate that polarization spectroscopy can be a helpful tool in determining the identity of individual bands and can assist in disentangling the underlying photocycle mechanism of photoreceptor proteins.

Introduction

Although PYP is a protein that has been studied by many different experimental approaches (see for a review (1–3)), many details of its photocycle remain to be solved. For a more elaborate introduction to PYP, the reader is referred to chapters 3 and 4 of this thesis. Recently, the first part of the photocycle mechanism (from photoexcitation up to 3 ns) has been studied in detail (see chapter 4). However, these measurements were done under anisotropy–free (magic angle) conditions, making no use of polarization information. Additional insight and structural information about the time–dependent dynamics of proteins can be obtained via the study of polarized measurements, as has been demonstrated for many systems, ranging from bacterial reaction centres to phytochromes and even water (4–6). Similar experiments on PYP have recently been published by Heyne *et al.* (7), who measured single wavelength time–traces with parallel and perpendicular polarizations to calculate the anisotropy value at the single wavelength for the ground– and for the I_0 state. This approach requires that the time–trace represents a contribution from only one spectral band (meaning overlapping spectra cannot be taken into account), and that no time-evolution occurs (which is a problem since bands can be formed or can shift in frequency as a function of time, which may obscure the 'pureness' of the calculated anisotropy at a single wavelength). A better approach is to allow the spectral bands to overlap,associate each band with its own anisotropy, and allow the bands to evolve in time. The pump–probe measurements on PYP described in this thesis were performed under magic angle conditions, implying that the pump–beam had a polarization rotated over 54.7^{o} with respect to the probe beam. This results in anisotropy–free measurements. This chapter discusses measurements made under conditions in which the polarization of the pump–beam is set perpendicular and parallel to the probe beam, thus allowing one to probe the change in orientation from individual atomic groups relative to the optical ground state transition dipole moment (from now on referred to as *ogstdm*). The ogstdm change gives information about the orientation of specific groups. As outlined in chapter 4, PYP isomerizes by a *trans* to *cis* process, and the chromophore's C=O flips to the other side of the chromophore's long–axis, thereby breaking the hydrogen bond with the protein backbone (see Figure 1). The isomerization and the C=O flip have been demonstrated to be two subsequent processes (chapter 4 and (8)). We expect the polarized measurements to yield more detail on the time–dependent isomerization of the chromophore, the orientation of the C=O group of the chromophore, and on the orientation of the C=O of Glu46 .

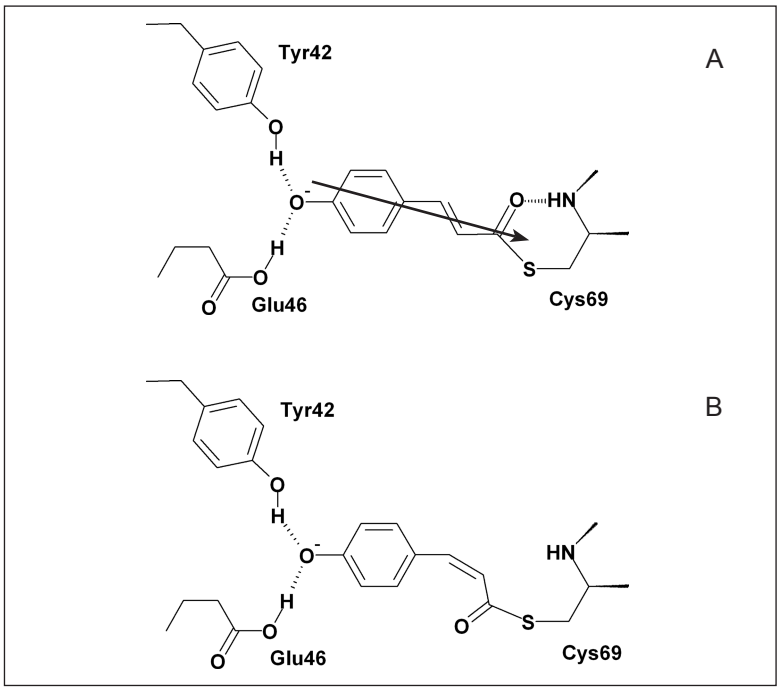


Figure 1 **Schematic drawing of the chromophore–binding pocket of PYP. The *trans*–chromophore (A) is embedded in a hydrogen–bonding network (dashed lines) with Glu46, Tyr42 and Cys69. The arrow represents the optical ground state transition dipole moment. After photoexcitation (B) the chromophore isomerizes first to form the *cis*–state, while the chromophore's carbonyl subsequently flips to the other side (breaking its hydrogen bond with Cys69) to form the I_0 state.**

Measurement Method

The samples consisted of highly concentrated protein solution of PYP from the bacterium *Halorhodospira halophila*, pressed between two CaF_2 windows that are separated by a 20–micron spacer. The sample was dissolved in 20 M Tris–D_2O buffer (pH 8), with an OD of 0.26 at 475 nm. A deuterated buffer was used to reduce solvent absorption in the probed spectral region. More detailed sample preparation and experimental conditions are described in chapters 3 and 4. To increase the signal–to–noise ratio, time gated pump–probe measurements were performed by averaging at only two time–points. The sample was excited at 475 nm with 300 nJ, and probed with a low intensity infrared beam (a few nJ) in the range from 1590–1776 cm^{-1}. The average of five background ΔOD (difference absorption) spectra (consisting of about 500 shots each) at a fixed delay of –10 ps were subtracted from the average of five ΔOD spectra taken at a positive time delay (measured directly after the background). One dataset therefore consists of a sequence of pre–zero and post–zero recordings. This procedure was repeated 100 times (we refer to each repetition of the procedure as a *scan*). The polarization of the pump–beam was set by a Berek rotator (Model 5540; New Focus, San Jose, CA) to an angle of 45° with respect to the polarization of the probe light. A BaF_2–polarizer (Medway Optics Ltd, Kent, UK), placed behind the sample, was used as an analyzer. This has the advantage that the overlap of the pump and probe beams in the sample remains unaffected when the analyzer is rotated.

The intensity of transmitted light polarized at parallel, magic angle, and perpendicular angle to the pump polarization was determined by rotating the analyzer to the appropriate amount, and the corresponding parallel, magic angle and perpendicular absorption difference spectra were calculated by taking the difference between pumped and un–pumped spectra. In this way, polarization sensitivity of the spectrograph did not play a role. To exclude polarization artefacts the polarization is randomly changed after every few scans between magic angle, perpendicular and parallel. The MCT detector (Infrared Associates, Inc. Stuart, FL, USA) used has 32 channels, which ultimately resulted in a final sampling resolution of ~6 cm^{-1}.

All collected polarized data were pre–processed to calculate the background–corrected data. The resulting averages were analysed with global analysis software (9), which allows for correlating spectral evolutions in time. To reduce experimental noise, the whole dataset was divided in 5 sets of 100 scans and a singular value decomposition matrix–analysis was performed. Since no reference probe pulse was used, the noise in the measured spectra consists mainly of so–called baseline noise, i.e. a flat, structure–less offset in the spectra, which is easily recognized from a singular value decomposition of the residual matrix (9, 10). The signal to noise ratio of the data is enhanced by subtracting the outer product of the first singular value pair of the residual matrix (being structure–less in the 'set' domain, *i.e.* the 5 sets of averages, and smooth in the wavelength domain) from the data, leading to a factor ~1.3 reduction in the noise for both 1–ps and 50–ps data. The final noise levels were ~56 μOD (1 ps) and ~44 μOD (50 ps), respectively.

The anisotropy r_l at fixed time t for the different evolving spectral bands (denote by the index l) was estimated from the relation:

$$\begin{bmatrix} \Delta A^{MA}(\lambda,t) \\ \Delta A^{\parallel}(\lambda,t) \\ \Delta A^{\perp}(\lambda,t) \end{bmatrix} = \sum_l c_{lt}\varepsilon_l(\lambda) \begin{bmatrix} 1 \\ 1+2r_l \\ 1-r_l \end{bmatrix}$$

The measured difference absorption signals with the probe beam set to magic angle (ΔA^{MA}), parallel ($\Delta A^{\parallel}$) or perpendicular ($\Delta A^{\perp}$) orientation with respect to the pump beam are time and wavelength (λ) dependent, and a function of the product of $c_{lt}\varepsilon_l$ and the anisotropy (9). The product $c_{lt}\varepsilon_l$ represents the spectrum of the spectral band at a certain time, and is built up from the time dependent concentration of the species c_{lt} and its spectrum ε_l. Because only two time points are measured, no concentration information is extracted and the product of $c_{lt}\varepsilon_l$ is estimated. Each

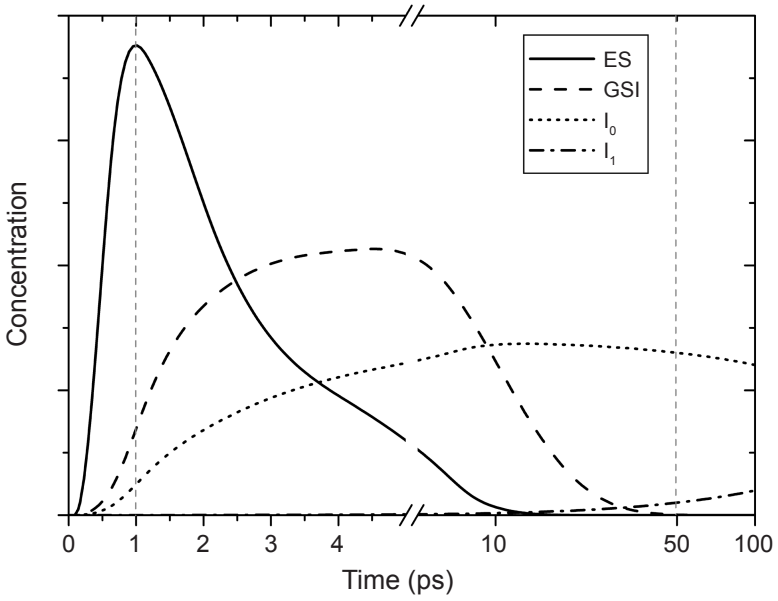


Figure 2 **Concentration profiles representing the fraction of species present at any given time (from photoexcitation, defined as 0 ps, up to 100 ps, based on the data of chapter 4). ES denotes the excited state population that simultaneously decays into the GSI (the ground state intermediate) and I_0 states. Only the latter populates the next photocycle intermediate I_1. The vertical dashed lines denote 1 and 50 ps, respectively. The time axis is linear up to 5 ps, logarithmic thereafter.**

spectral band may actually consist of multiple bands, as long as the anisotropy value is the same. The sum of all bands forms a spectrum, the intensity of which is fitted to the measured polarized spectra. The anisotropy value is related to the angle β between the ogstdm and the vibrational transition dipole moment of the probed (individual) bands, according to the formula $r = 0.2 \times \left(3 \times \cos^2 \beta - 1\right)$, implying that the anisotropy theoretically lies between –0.2 and 0.4.

Results

Polarized difference absorption measurements (initiated by optical excitation at 475 nm) were performed at fixed time delays of 1 and 50 ps. The 1–ps data mainly represent the excited state (ES) of PYP, while the 50–ps data mainly represent the I_0 form (see Figure 2). The recorded difference spectra are presented in Figure 3 and 4, respectively. Note that negative bands correspond to the ground state, and positive ones to the excited and product state(s). The control spectrum measured at magic angle lies igenerally between the parallelly and perpendicularly polarized spectra, as expected. However, there are some minor discrepancies in the 50–ps data (see 1628 and 1640 cm^{-1} in Figure 4), which can be considered as indicators of the achievable signal–to–noise ratio. Because the presented spectral region obviously consists of multiple spectral features, each feature is analysed separately, modeled by $c_{it}\varepsilon_i$, and the anisotropy of each feature is determined. Even then, each feature may consist of more than one band, but this is allowed as long as the anisotropy is identical. The estimated anisotropies are

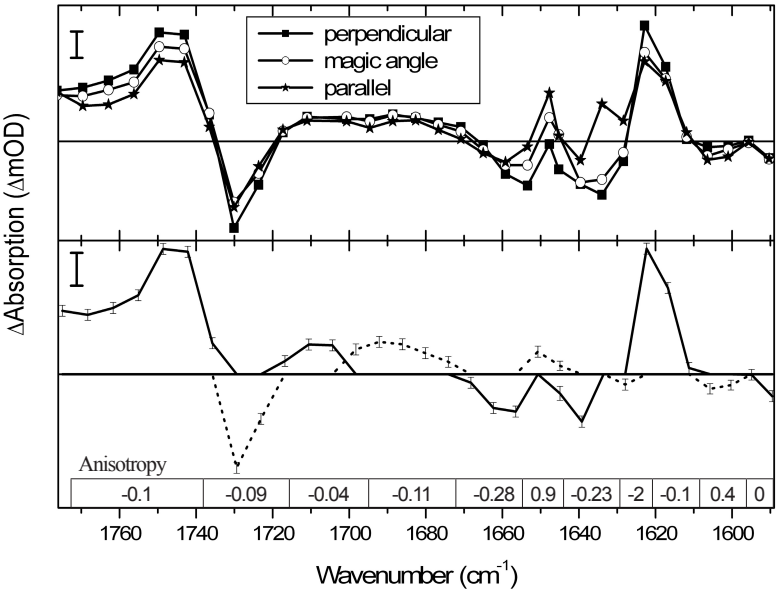


Figure 3 **Polarization dependent measurements on PYP in D_2O at a fixed time delay of 1 ps. The vertical bar indicates 0.1 mOD. Difference spectra are shown in the top panel for perpendicular, magic angle and parallel polarizations of the probe beam with respect to the pump beam, respectively. The bottom panel shows the individual bands (with error bars) used to model the spectra in the top panel. The anisotropies calculated from the perpendicular, magic angle and parallel intensities are shown for each band in the table (lower panel), with the values directly below each corresponding band. For clarity the bands have alternating line styles (solid/dot).**

shown in the tables in the bottom panels of Figure 3 and 4, directly below the corresponding bands. Some small bands in the 1–ps data (at 1628 and 1648 cm^{-1} in Figure 3) have unrealistically high values (r > 0.4), possibly due to overlapping bands (leading to compensating anisotropy values), or to their spectral width (note that the width of the band at 1648 cm^{-1} approaches the sampling resolution of ~6 cm^{-1}).

Discussion

The pump–probe measurements presented here contain, besides spectral and temporal information, also spatial information. To correctly model the information that the spectra contain, each band has been assigned its own anisotropy. Based on the most recent time–resolved infrared data (presented in chapter 4), the transient population of photoexcited PYP after 1 ps actually consists of a mix of mainly ES (~70%, see Figure 2), some ground state intermediate (GSI; ~20%) and a small amount of the first stable photocycle intermediate I_0 (~10%). The 50 ps data however represent a nearly pure I_0–state. The difference spectra measured at delays of 1 and 50 ps are shown in Figure 3 and 4 in the top panels, while the spectral bands used to model the collected spectra are shown in the bottom ones. The measured anisotropy data collected at 1 and 50 ps are compared with previous reports and discussed in terms of the photocycle of PYP.

New feature?

The band at 1634 cm^{-1} in the 1–ps data has a peculiar double nature, with a positive parallel and a negative perpendicular

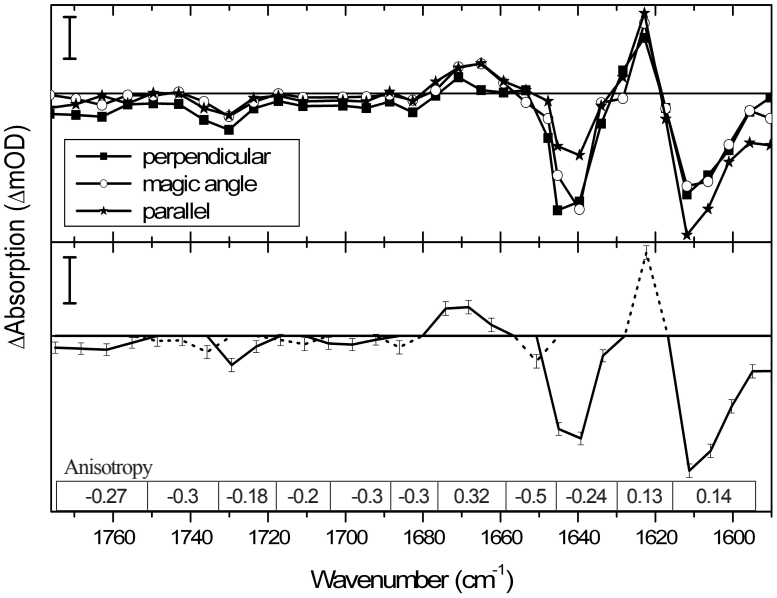


Figure 4 **Polarization dependent measurements on PYP in D_2O at a fixed time delay of 50 ps. The vertical bar denotes 0.1 mOD. Difference spectra are shown in the top panel for perpendicular, magic angle and parallel polarizations of the probe beam with respect to the pump beam, respectively. The bottom panel shows the individual bands (with error bars) used to model the spectra in the top panel. The anisotropies calculated from the perpendicular, magic angle and parallel intensities are shown for each band in the table (lower panel), with the values directly below each corresponding band. For clarity the bands have alternating line styles (solid/dot).**

contribution (and a negative magic angle spectrum). This feature cannot be well-described in terms of the model applied, due to unresolved amplitude compensation effects. The feature may be an artefact, *i.e.* a baseline shift. However, if the feature is fitted with two bands of opposite amplitude (~1/3 higher than the measured maximum of the parallel and perpendicular signals at 1634 cm^{-1}), having anisotropies of 0.4 and –0.2 for the positive and negative band, respectively, a magic angle signal size is obtained of about –0.12. We therefore hypothesize that this feature consists of two bands, one close to the parallel (the positive feature) and one close to the perpendicular orientation (the bleach) with respect to the ogstdm. The positive (parallel) feature at 1634 cm^{-1} is new and has not been observed before (it is undetectable by magic angle measurements). Unfortunately, we are unable to validate this hypothesis by our current level of modelling.

It is clear that, apart from the spatial information, the polarized data also yield additional information on the splitting of unresolved bands, *e.g.* the 1634 cm^{-1} band, or partially resolved bands, such as 1670–1720 cm^{-1} in the 1–ps data, and on the overlap between bleached ground state and product state absorption bands, as can be seen for the bands above 1710 cm^{-1} in the 1–ps data.

The probed spectral region (1590–1776 cm^{-1}) contains the infrared signatures of C=O groups (11), among others, and is therefore ideal for monitoring the C=O groups of the chromophore and Glu46. It also contains signatures of PYP's *trans–cis markers* (presumably C=C stretch modes in the 1600–1630 cm^{-1} region (12–16) and chapter 7).

Glu46

The most obvious features in the 1–ps data in Figure 3 are located above 1715 cm^{-1}. These bands correspond to a signal from the Glu46–residue, which is hydrogen bonded to the phenol moiety of the chromophore. The C=O group of Glu46 is reported to be located at 1740 cm^{-1} (8, 12, 14, 17) and shown to downshift with ~10 cm^{-1} upon deuteration (17). Upon photoexcitation, we observe it to upshift in frequency from 1730 cm^{-1} to 1743/1750 cm^{-1} (Figure 3), while retaining its anisotropy of ~–0.1, corresponding to an angle β = ~65° (± 7 °) between the ogstdm and the C=O group. Heyne *et al.* reported an angle of ~84° (± 6°) for both the bleach and the upshifted band (7). A systematic error in either dataset may cause this discrepancy. The fact that the product band is relatively broad is attributed to the presence of a mix of ES and GSI; it upshifts an additional 10 cm^{-1} going from ES to GSI (8, 12). The 1730 cm^{-1} to 1743/1750 cm^{-1} upshift is interpreted as a weakening of the Glu46–chromophore hydrogen bond. Note, however, that the anisotropy value remains constant during this shift, indicating that the carbonyl of Glu46 does not change its orientation during this part of the photoreaction.

At the 50–ps time delay (see Figure 2 and 4), the transient spectra have completely changed with respect to the 1–ps delay; they now represent the I_0–species (12, 18–20). In chapter 3 and ref. (12) the Glu46 band is reported to decrease in intensity and to downshift (from 1740 to 1732 cm^{-1}), which was interpreted as a strengthening of the hydrogen bond between Glu46 and the phenol moiety of the chromophore in the I_0 state. However, there are some contrasting reports about this downshift. A polarization study similar to the one presented here found no transient signal of either polarization (perpendicular nor parallel) for the Glu46 bands representing I_0 (7). Low temperature FTIR measurements report an ~8 cm^{-1} downshift with the formation of I_1 (14). Because the differences between the I_0 and I_1 difference spectra are negligible in this spectral region (8, 12), the interpretation put forward in chapter 3 is preferable. In Figure 4, the feature at 1724 cm^{-1} has an anisotropy value of about –0.2 (β = ~90° ± 39°). Presuming a negative baseline, this feature may seem to appear to downshift to 1717 cm^{-1} (and having the same anisotropy). However, due to the negative baseline, the anisotropy is impossible to model. The bleach at 50 ps is observed at virtually the same location as at 1 ps (1724 cm^{-1} vs. 1730 cm^{-1} , respectively), given the spectral width and sampling resolution. We find a different orientation at 1 ps and 50 ps (β = ~65° (± 7 °) vs. ~90° ± 39°, respectively). The differences in signal size between 50 ps and 1 ps in the 1715–1730 cm^{-1} region, and the overlap of a small feature above 1730 cm^{-1} (r = –0.3) with the band at 1724 cm^{-1}, are most likely causing the measured differences in orientation and error. The orientation measured at 1 ps is evidently more reliable. Although

a negative baseline hampers the interpretation of the 50–ps data, the polarized measurements presented here seem to confirm the hydrogen bond strengthening hypothesis and show that the C=O of Glu46 has a similar orientation in the ground state and I_0, in agreement with the X–ray diffraction studies (21).

Isomerization markers

Markers for the isomerization process appear in the spectra on the low frequency side of the probed infrared region, where a band appears at 50 ps that was absent at 1 ps. This band is located at 1612 cm^{-1}, and has been assigned as a *trans–cis* marker (12–16). Our polarized data show that two bands in this region (at 1612 and 1623 cm^{-1}, respectively) have the same anisotropy at 50 ps (r = ~0.13; β = ~42° ± 10°). The appearance of these bands at 50 ps (the band at 1623 cm^{-1} at 1 ps has a different anisotropy of –0.1, and therefore a different origin; it is attributed to ES absorption) indicates that the chromophore has isomerized on a picosecond timescale. Calculations on model chromophores (chapter 7 and ref. (7)) show, however, that the stretching mode of the central C=C bond *down*shifts with 10–20 cm^{-1} from ~1640 cm^{-1} (mixed with the chromophore's C=O stretching mode), upon the formation of I_0. It is therefore likely that the two bands at 1612 and 1623 cm^{-1} have different origins. When we follow the assignment from the quantum chemical calculations and calculate the change in orientation of the C=C bond with respect to the ogstdm from the anisotropies of the negative band at 1640 cm^{-1} and the positive band at 1623 cm^{-1} we find that it changes from perpendicular to 42°, due to the isomerization. However, due to mixing of the bleach at 1640 cm^{-1} with the chromophore's C=O stretching mode (see below), the anisotropy value calculated is a mix between C=C and chromphore C=O.

C=O of the chromophore

Another infrared signature that appears in the presented spectral window is that of the C=O group of the chromophore. Upon photoexcitation, this C=O mode is reported to upshift from 1635/1640 cm^{-1} in the ground state to 1665 cm^{-1} in I_0 (see chapter 3 and ref. (22)). As shown in chapter 3 and 4, no product band is observed in the ES. This absence is explained in terms of a photoinduced rearrangement of the conjugation over the chromophore, rendering the double bond of the chromophore's C=O more single bond (with a more double bound character), while maintaining its hydrogen bond to the backbone. The characteristic frequency of a C–O vibrational mode is expected in the 1250 cm^{-1} region (11), far outside the probed spectral window in this study. The polarized 1–ps data presented here indeed have a bleach at 1634–1640 cm^{-1} (r = –0.23; β = ~90° ± 20°), while a product band (with identical anisotropy) is absent in the 1640–1670 cm^{-1} region. The chromophore's C=O bleach at 50–ps

is slightly upshifted to 1640 cm^{-1} in Figure 4, and a product band appears at 1665 cm^{-1}. However, whereas the anisotropy of the band at 1640 cm^{-1} is similar to that in the 1–ps spectrum (–0.24) the anisotropy of the product band at 1665 cm^{-1} is different: r = 0.32; β = ~21° ± 20°. This change in angle upon isomerization can be explained by the strong degree of overlap of the chromophore's C=O bleaching mode with other modes, such as the stretching modes of the chromophore's central C=C bond and the phenolic oxygen, according to calculations (see chapter 7 and ref. (7), but note that in the latter reference β changes from 57° ± 5° to 69° ± 10°). The observed orientation differences of the chromophore's C=O in I_0 between literature (7) and this work (21° vs. 69°) are too large to be explained by the experimental errors, and could be caused by systematic errors in either dataset. This issue remains therefore subject to further investigation. The change in orientation we observe (from β = ~90° ± 20° at 1 ps to ~21° ± 20° at 50 ps) can be explained by the formation of the isomerized chromophore, which has a flipped C=O group (no longer hydrogen bonded to the backbone and with an altered orientation).

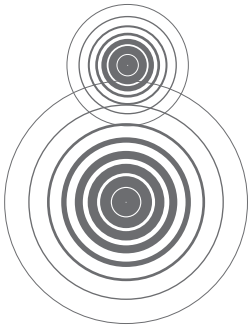
Conclusions

Polarized infrared measurements add an extra dimension to pump–probe spectroscopy, forming a useful tool for the assignment of infrared bands of PYP. Complicating factors for the type of infrared anisotropy measurements presented here are the signal–to–noise ratio, the simultaneous presence of multiple states at specific time delays, and normal mode mixing for molecules with many delocalized vibrations. Time gated anisotropy measurements were recorded at fixed time delays of 1 and 50 ps to obtain more insight in the photoinduced dynamic changes in the chromophore's hydrogen bonding network. Our results demonstrate that I_0 possesses a similar anisotropy value at 1 and 50 ps for the bleach of Glu46's C=O group, which is hydrogen bonded to the chromophore. However, at 50 ps this band is probably downshifted, confirming that the hydrogen bond is indeed strengthened with I_0. The anisotropy data confirm the hypothesis that photoexcitation renders the chromophore's C=O more single bond, while maintaining its hydrogen bond to the backbone. At 50 ps, the carbonyl has changed its orientation and has broken its hydrogen bond to the backbone. The polarized data presented here have increased our insight into the photoactivation mechanism of PYP.

References

1. Larsen, D. S. & van Grondelle, R. (2005) *Chemphyschem* 6, 828-837.
2. Cusanovich, M. A. & Meyer, T. E. (2003) *Biochemistry* 42, 4759-4770.
3. Hellingwerf, K. J., Hendriks, J. & Gensch, T. (2003) *Journal of Physical Chemistry A* 107, 1082-1094.
4. Jonas, D. M., Lang, M. J., Nagasawa, Y. & Fleming, G. R. (1996), pp. 46-7.
5. Savikhin, S., Wells, T., Song, P. S. & Struve, W. S. (1993) *Biochemistry* 32, 7512-7518.
6. Fecko, C. J., Loparo, J. J., Roberts, S. T. & Tokmakoff, A. (2005) *Journal of Chemical Physics* 122, -.
7. Heyne, K., Mohammed, O. F., Usman, A., Dreyer, J., Nibbering, E. T. J. & Cusanovich, M. A. (2005) *Journal of the American Chemical Society* 127, 18100-18106.
8. van Wilderen, L. J. G. W., van der Horst, M. A., van Stokkum, I. H. M., Hellingwerf, K. J., van Grondelle, R. & Groot, M.-L. (2006) *Proc Natl Acad Sci U S A* 103, 15050-15055.
9. van Stokkum, I. H. M., Larsen, D. S. & van Grondelle, R. (2004) *Biochimica et Biophysica Acta-Bioenergetics* 1657, 82-104.
10. Hoff, W. D., van Stokkum, I. H. M., van Ramesdonk, H. J., van Brederode, M. E., Brouwer, A. M., Fitch, J. C., Meyer, T. E., van Grondelle, R. & Hellingwerf, K. J. (1994) *Biophysical Journal* 67, 1691-1705.
11. Barth, B. (2000) *Progress in Biophysics & Molecular Biology* 74, 141-173.
12. Groot, M.-L., van Wilderen, L. J. G. W., Larsen, D. S., van der Horst, M. A., van Stokkum, I. H. M., Hellingwerf, K. J. & van Grondelle, R. (2003) *Biochemistry* 42, 10054-10059.
13. Brudler, R., Rammelsberg, R., Woo, T. T., Getzoff, E. D. & Gerwert, K. (2001) *Nat Struct Biol* 8, 265-70.
14. Imamoto, Y., Shirahige, Y., Tokunaga, F., Kinoshita, T., Yoshihara, K. & Kataoka, M. (2001) *Biochemistry* 40, 8997-9004.
15. van Thor, J. J., Pierik, A. J., Nugteren-Roodzant, I., Xie, A. H. & Hellingwerf, K. J. (1998) *Biochemistry* 37, 16915-16921.
16. Kim, M., Mathies, R. A., Hoff, W. D. & Hellingwerf, K. J. (1995) *Biochemistry* 34, 12669-72.
17. Xie, A., Hoff, W. D., Kroon, A. R. & Hellingwerf, K. J. (1996) *Biochemistry* 35, 14671-14678.
18. Baltuška, A., van Stokkum, I. H. M., Kroon, A., Monshouwer, R., Hellingwerf, K. J. & van Grondelle, R. (1997) *Chemical Physics Letters* 270, 263-266.
19. Ujj, L., Devanathan, S., Meyer, T. E., Cusanovich, M. A., Tollin, G. & Atkinson, G. H. (1998) *Biophysical Journal* 75, 406-12.
20. Imamoto, Y., Kataoka, M., Tokunaga, F., Asahi, T. & Masuhara,

H. (2001) *Biochemistry* 40, 6047-6052.
21. Ihee, H., Rajagopal, S., Srajer, V., Pahl, R., Anderson, S., Schmidt, M., Schotte, F., Anfinrud, P. A., Wulff, M. & Moffat, K. (2005) *Proceedings of the National Academy of Sciences of the United States of America* 102, 7145-7150.
22. Unno, M., Kumauchi, M., Sasaki, J., Tokunaga, F. & Yamauchi, S. (2002) *Biochemistry* 41, 5668-5674.



Vibrational Analysis of the Isomerization Reaction in Photoactive Yellow Protein

07

L.J.G.W. van Wilderen, M. Di Donato, R. van Grondelle, M.L. Groot

Manuscript in preparation

Abstract

Quantum chemical calculations of vibrational mode frequencies of the chromophore of Photoactive Yellow Protein are presented in order to obtain a better understanding of the photo–induced trans–cis *isomerization process in PYP. Within the protein, the chromophore is embedded in a hydrogen bond network, and therefore calculations are performed with water molecules to mimic these hydrogen bond interactions. Normal mode analysis in combination with potential energy distribution analysis is used to facilitate an objective vibrational assignment. The first stable cis ground state intermediate of the chromophore (I_0) is characterized by a broken hydrogen bond between the chromophore's carbonyl and the protein backbone. Density functional theory (DFT) and Hartree-Fock (HF) levels of theory are used for the calculations, and are compared with experimental data. The DFT method results are in a closer agreement with the experimental data than those obtained by HF, and especially the ground state calculations produce a general correspondence with the measurements. All calculations presented here show a strong degree of mode mixing for virtually every mode. We report multiple isomerization markers for the chromophore that span the whole mid–infrared fingerprint region, and these results represent a helpful tool in deciding where to start with various experimental approaches. Ultimately these results assist in understanding protein function.*

Introduction

Infrared spectra are packed with information about the vibrational structure of the chromophore and the protein. Identifying the true nature of a single vibrational band in the spectrum is not a trivial task. A time–consuming and costly way is to isotope–label one atom in the protein, or to mutate one residue, and see what effect it has on the experimental infrared spectra. The chromophore can also be synthesized in solution, and probed by infrared spectroscopic techniques to isolate the chromophore modes. Another approach is to calculate the vibrational spectrum, which allows the determination of the exact nature of one vibration, although the experimentally determined intensity and location can in general not be reproduced fully reliably (*1*). Calculations are, however, consistent in giving trends (one vibration has for example always a higher frequency than another), and are therefore an extra means to help understanding experimentally acquired data.

The system studied is photoactive yellow protein (PYP, see (2–4) for a review). It has a *p*–coumaric acid molecule as its chromophore (see Figure 1A), which isomerises in a few picoseconds upon electronic excitation (see Figure 1B) (5–7). A crucial step in the formation of a stable *cis*–ground state intermediate is the breaking of the hydrogen bond between the residue Cys69 and the chromophore's carbonyl. A biologically productive state is only formed when the H–bond is broken, while an intact one leads to the reformation of the ground state denoted pG (via a ground state intermediate (7–10)). The PYP photocycle has been characterized by a variety of experimental techniques, such as fluorescence (11, 12), (time–resolved) FTIR (13–15), (time–resolved) X–ray crystallography (16–19), NMR (20), Stark spectroscopy (21) and pump(–dump)–probe spectroscopy (5–7, 10, 22). The protein environment has been found to have a profound effect on the chromophore dynamics, as the ultrafast (femto– to picosecond) dynamics of photo–excited model chromophores in solution prove. The chemical properties of the terminating moiety of the chromophore seem to determine if the chromophore isomerizes or not. For example, the chromophore with a phenyl group attached to the sulphur atom (thiophenyl–*p*–hydroxycinnamate) isomerizes (23), but with a methyl group (thiomethyl *p*– hydroxycinnamate) it does not (24). This matter is still under debate and the discussion extends also to the gas phase where both the neutral and the anionic form of *trans*–*p*–coumaric acid were found to isomerize (25, 26). Infrared spectroscopy is an ideal technique to determine if a molecule isomerizes or not, since several of the IR active modes have been identified as markers for isomerization (6, 27). For a better interpretation of these spectra and of the isomerization process the need arises for a full understanding of the origins of all spectral features. In this work we present quantum chemical calculations on PYP to compare with our earlier obtained experimental infrared difference spectra and thus obtain a better

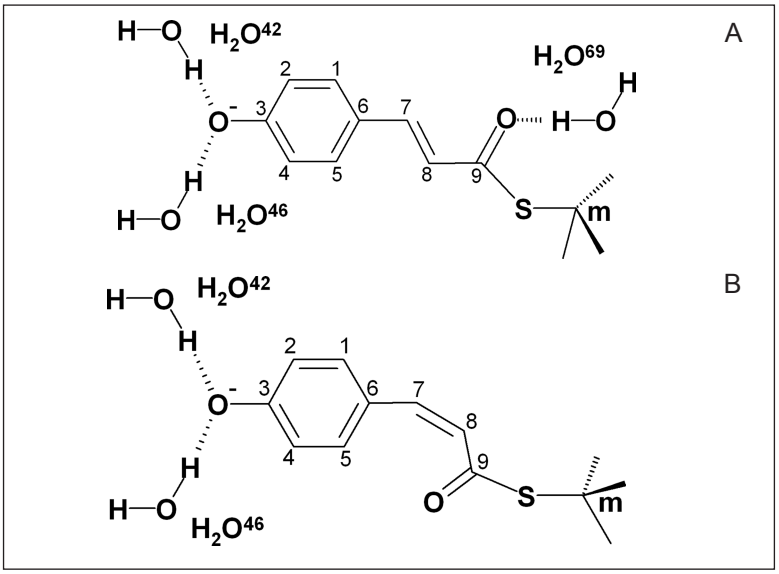


Figure 1 **Schematic drawing of the model chromophore conformations used for the calculations. The chromophores are terminated by a methyl group, denoted by an *m* in both panels. In panel A, the ground state conformation of PYP based on the 1NWZ pdb–structure file is drawn. The three water molecules (H_2O^{42}, H_2O^{46} and H_2O^{69}) are H–bonded with the chromophore, and replace the amino acid residues Tyr42, Glu46 and Cys69, respectively. In panel B, the first stable photocycle intermediate I_0 of PYP based on the 3PYP pdb–structure file is drawn. The two water molecules (H_2O^{42} and H_2O^{46}) are H–bonded to the chromophore and replace the amino acid residues Tyr42 and Glu46, respectively.**

understanding of the PYP isomerization process.

The normal modes of different model chromophores have been calculated previously (6, 28–30); Of the neutral and negatively charged methyl–ester–*p*–coumaric acid (28), the negatively charged chromophore with Cys69 attached (*6*), trans–S–phenyl–thio–*p*–hydroxycinnamate (pCT) and pCT$^-$–3H_2O (with three water molecules attached to mimic hydrogen–bonding interactions (29)), and 4–hydroxycinnamyl–methyl–thioester (HCMT (30)). Here, calculations on the negatively charged *trans* and *cis* forms of methyl–thio–*p*–coumaric acid are presented. The interactions between the chromophore and the direct hydrogen–bonding partners in the chromophore–binding pocket, Glu46 and Tyr42 on the phenolate oxygen and Cys69 on the carboxylic oxygen, are mimicked by the insertion of three water molecules at their respective amino–acid positions. In order to estimate the frequency shift of the carbonyl stretching vibration due to the breaking of the H–bond with Cys69, the calculation on the *cis* form does not include a water molecule at position Cys69 (compare Figure 1A with B). The isomerization process is investigated by monitoring the changes of the normal modes of the chromophore in the *trans* and *cis* form. The results from the normal mode analysis (NMA) are compared with the experimentally observed infrared difference spectra presented in chapters 3, 4.

Methods

Calculations

The quantum chemistry package of WinGamess (31) is used for the NMA calculation. As starting geometries for the optimisation and frequency runs, the molecular structures from the Brookhaven

protein databank are used (32). The 1NWZ (DOI: 10.2210/pdb1nwz/pdb) and 3PYP (DOI: 10.2210/pdb3pyp/pdb) pdb–files are the structures with the highest resolution available, and taken to represent the *trans*– and the *cis*–conformation, respectively. Hydrogen atoms are manually added to the chromophore, but the phenol–moiety remains deprotonated. The hydrogen bond–donating residues Cys69, Glu46 and Tyr42 are exchanged for water molecules and the distances from these residues to the chromophore are set to the values found in the pdb–files. A methyl group terminates the chromophore after the sulphur–atom. The level of theory is stepwise increased from STO–3G to HF/6–311++G(p,d) for *ab initio* computations (referred to as HF) and B3LYP/6–311++G(d,p) for DFT. These basis sets have shown to adequately predict experimentally observed frequencies for a large set of molecules, and are considered to be currently the most cost–effective (*i.e.* high accuracy with relatively low computational cost) for NMA (33). Because molecules thermally vibrate, a geometry optimization is performed on the molecular structures taken from the ('frozen') X–ray diffraction data to calculate the (local) minimum on the potential energy surface (PES; see also chapter 1). No geometry constraints are applied. After the first optimisation run the chromophore becomes planar for both the *trans* and *cis*–forms, remaining planar also with the more extended basis sets that are employed. Equilibrium geometries are determined with the same level of theory as the frequency run to determine if the (local) minimum PES is reached (*1*). Diffusive shells (*s* and *sp*) have proven to be necessary for determining the (local) PES minimum for all but the HF–run with the 1NWZ–based geometry. The addition of diffusive shells is beneficial for the calculation of intermolecular H–bonds (34, 35) and the prediction of IR intensities (36), while it has hardly an effect on the calculation of frequencies (37). All frequency runs presented here have no imaginary frequencies. The focus lies on the protein 'fingerprint–region' (970–1750 cm^{-1}), and therefore not all (3*n*–6, with *n* the number of atoms) calculated modes are shown. The scaling factors used for HF and B3LYP are 0.9054 and 0.9614, respectively (33). Transition dipole moments are calculated by WinGamess (10.1/9.4 for 1NWZ, and 11.4/11.0 for 3PYP for HF and DFT, respectively).

The potential energy distribution (PED) is calculated from the WinGamess output file with FCONV and VIBCA (38). This approach maps the intrinsic frequencies generated by WinGamess onto internal coordinates, and not only takes into account that different masses contribute differently to the kinetic energy of every observed frequency, but also renders the assignment process objective and reproducible. The internal coordinates were generated with MacMolPlt (39). All contributions in the PED add up to 100%, but due to inaccurate mapping of the internal coordinates on the normal coordinates, some easily recognizable

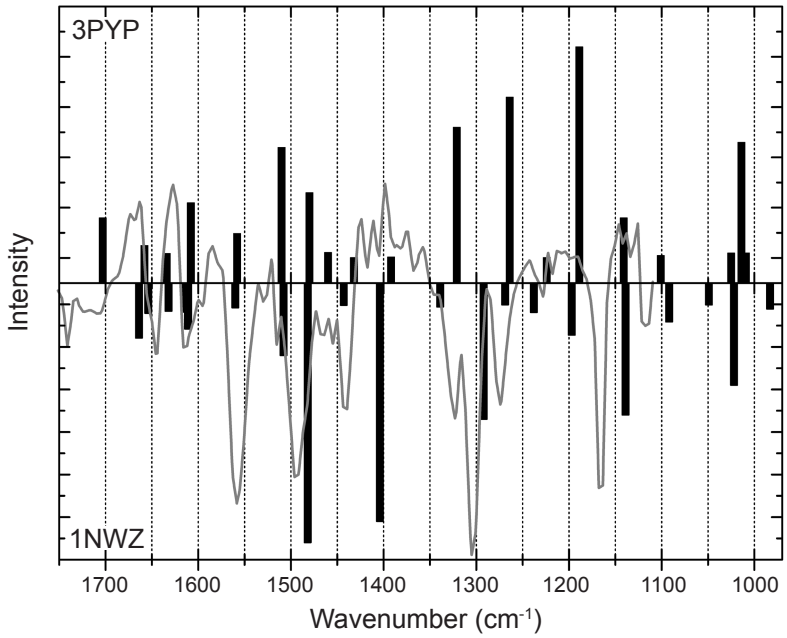


Figure 2 **Comparison of the HF calculations on 1NWZ (bottom panel) and 3PYP (top panel), and the experimental I_0-pG difference spectrum (grey line, ΔmOD x 20 as function of wavenumber). The calculated intensities are plotted as a function of wavenumber (scaled by 0.9054), and have an offset of an additional 5 units (negative for the top panel, positive for the bottom one) to make the low intensity frequencies more visible. The shared x-axis is the baseline for the experimental spectrum.**

compensation contributions appear (resulting in minor values with negative contributions). However, this does not influence the outcome of the final assignment.

Infrared spectroscopy

The calculations in this chapter are compared with the experimental data already presented in chapter 4. For any experimental details on the pump–probe setup that was used and the sample preparation, the reader is referred to chapter 4. In summary, a highly concentrated wild type PYP sample (OD^{446nm} ≈ 1) is photoexcited with a 475–nm laser pulse, and probed in the infrared in the 1126–1750 cm^{-1} region. The arrival of the pump–pulse with respect to the probe pulse is delayed from –16 ps to 3 ns, and the transient absorption of the 'pumped' sample is compared with the 'un–pumped' sample. The resulting infrared difference spectra were subjected to a global analysis, and only the spectrum that is formed in ~1 ps, and has a lifetime of ~700 ps (corresponding to I_0–pG), is shown in Figure 2.

Results and Discussion

The NMA on the model chromophores shown in Figure 1A and B are compared with the experimentally obtained data on the native protein. NMA allows for the determination of the origin of each experimentally observed frequency. An objective insight in the results from the NMA is obtained by calculating the PED (40, 41), in which a mass–weighed contribution of each normal

Freq.	Int.	PED (%)							
1664	7.9	50	$\delta(O_9{=}C_9{-}C_8)$	33	$\nu(O_9{=}C_9)$	13	$\delta(C_8{=}C_7{-}C_6)$		
1654	2.1	100	$\gamma(H_2O^{42}{-}O_3)$						
1632	1.6	100	$\delta(H_2O^{46})$						
1614	1.8	52	$\nu(OH^{69})$	33	$\gamma(H_2O^{69}{-}H_m)$	10	$\delta(H_2O^{69})$		
1611	5.8	22	$\nu(C_4{-}C_5)$	14	$\gamma(H_2O^{46}{-}O_3)$	9	$\nu(C_1{-}C_2)$	9	$\delta(C_1{-}C_2{-}C_3)$
1560	0.8	41	$\nu(O_3{-}C_3)$	19	$\delta(C_1{-}C_6{-}C_5)$				
1508	12	27	$\delta(C_7{-}C_6{-}C_1)$	22	$\delta(C_3{-}C_4{-}C_5)$	18	$\delta(C_2{-}C_1{-}C_6)$		
1482	56	30	$\nu(C_4{-}C_5)$	11	$\delta(H_7{-}C_7{=}C_8)$	8	$\nu(C_1{-}C_2)$		
1443	0.2	24	$\nu(C_1{-}C_2)$	14	$\nu(H_2{-}C_2)$	11	$\delta(C_1{-}C_2{-}C_3)$	10	$\delta(H_8{-}C_8{=}C_7)$
1404	51	31	$\nu(O_3{-}C_3)$	9	$\nu(H_5{-}C_5)$				
1339	0.6	26	$\delta(H_8{-}C_8{=}C_7)$	11	$\nu(C_6{-}C_7)$	11	$\nu(C_9{-}C_8)$		
1292	27	26	$\delta(C_7{-}C_6{-}C_1)$	16	$\nu(H_5{-}C_5)$	14	$\nu(C_1{-}C_6)$	13	$\nu(H_4{-}C_4)$
1269	0.1	48	$\delta(C_8{=}C_7{-}C_6)$	35	$\delta(C_9{-}C_8{=}C_7)$	12	$\nu(H_7{-}C_7)$	7	$\nu(C_7{=}C_8)$
1238	1.9	17	$\nu(C_2{-}C_3)$	15	$\nu(H_2{-}C_2)$	14	$\delta(O_3{-}C_3{-}C_2)$	11	$\delta(C_3{-}C_4{-}C_5)$
1197	7.2	14	$\nu(C_5{-}C_6)$	12	$\delta(C_1{-}C_6{-}C_7)$	12	$\nu(C_4{-}C_5)$	10	$\nu(H_4{-}C_4)$
1139	26	14	$\nu(H_1{-}C_1)$	13	$\nu(H_5{-}C_5)$	12	$\nu(H_2{-}C_2)$		
1092	4.1	15	$\nu(H_2{-}C_2)$	10	$\nu(H_5{-}C_5)$	10	$\nu(C_1{-}C_2)$	10	$\delta(C_7{-}C_6{-}C_1)$
1049	0.1	41	$\nu(C_9{-}C_8)$	14	$\nu(O_9{=}C_9)$	13	$\delta(C_6{-}C_2{-}C_1)$	11	$\delta(C_9{-}C_8{=}C_7)$
1022	19	96	$\gamma(H_7{-}C_7{=}C_8{-}H_8)$						
983	1.1	35	$\gamma(H_5{-}C_5{-}C_4{-}C_3)$	20	$\gamma(H_1{-}C_1{-}C_2{-}C_3)$	15	$\gamma(C_3{-}C_4{-}C_5{-}C_6)$		

The frequencies are shown in cm^{-1}, with their respective calculated intensities. The modes that are localized mainly on the methyl group (1455, 1439 and 1353 cm^{-1}) are excluded from the table, and only the modes contributing ~10% or more to the PED are shown. ν Denotes a stretching mode, δ a bending mode and γ an out–of–plane mode.

Table 1 **Calculated frequencies (HF/6–311–G(p,d) scaled by 0.9054), intensities and PED for the chromophore in the ground state, based on the 1NWZ structure.**

mode to the potential energy is calculated. *Ab initio* (Hartree–Fock, HF) and density functional theory (DFT; the B3LYP electron correlation functional is used) are employed for both the *trans* and *cis* conformations. HF uses the assumption that each electron sees all of the others as an average field, thus neglecting electron correlation and interaction. DFT, in combination with an electron correlation functional (such as B3LYP), includes these interactions and results generally in better results, but renders the calculation more time–consuming (1). The calculated vibrational frequencies for the *trans* and the *cis* forms of the negatively charged chromophore are listed in Tables 1–4. Only the modes that contribute ~10% or more to the PED are reported with their calculated intensities. Because of the high degree of π–conjugation in the molecular structure of the PYP chromophore, its vibrational modes are highly

Freq.	Int.	PED (%)							
1662	0.9	81	$\delta(H_2O^{46})$	10	$\delta(OH^{46}–O_3)$				
1659	4.0	82	$\delta(H_2O^{42})$	10	$\delta(OH^{42}–O_3)$				
1652	0.8	48	$\delta(H_2O^{69})$	34	$\nu(O_9=C_9)$				
1633	4.5	40	$\delta(H_2O^{69})$	29	$\nu(O_9=C_9)$				
1616	0.3	21	$\nu(C_4–C_5)$	13	$\nu(C_7=C_8)$	10	$\nu(C_1–C_2)$		
1555	53	17	$\nu(O_9=C_9)$	12	$\nu(C_4–C_5)$	9	$\nu(C_9–C_8)$		
1527	37	36	$\nu(O_3–C_3)$	9	$\delta(H_4–C_4–C_5)$				
1486	0.3	22	$\nu(C_1–C_6)$	18	$\nu(C_5–C_6)$	12	$\nu(C_4–C_5)$		
1465	3.6	29	$\nu(C_1–C_2)$	26	$\nu(C_4–C_5)$	10	$\delta(H_2–C_2–C_1)$		
1400	1.6	23	$\nu(O_3–C_3)$	20	$\nu(C_6–C_7)$	18	$\delta(H_8–C_8=C_7)$		
1370	0.1	20	$\delta(H_8–C_8=C_7)$	18	$\delta(H_1–C_1–C_2)$	16	$\delta(H_2–C_2–C_1)$		
1302	9.6	16	$\delta(H_5–C_5–C_4)$	14	$\delta(H_4–C_4–C_5)$	13	$\nu(C_6–C_7)$		
1281	2.0	36	$\nu(C_7=C_8)$	33	$\delta(H_7–C_7=C_8)$	10	$\delta(H_8–C_8=C_7)$		
1266	0.1	23	$\delta(H_4–C_4–C_5)$	16	$\delta(H_2–C_2–C_1)$	15	$\nu(C_2–C_3)$	14	$\delta(C_3–C_4–C_5)$
1229	5.1	16	$\delta(H_7–C_7=C_8)$	16	$\delta(H_1–C_1–C_2)$	13	$\nu(C_8–C_6)$		
1144	7.5	25	$\delta(H_5–C_5–C_4)$	21	$\delta(H_2–C_2–C_1)$	20	$\delta(H_1–C_1–C_2)$		
1088	0.4	20	$\nu(C_9–C_8)$	14	$\delta(H_5–C_5–C_4)$	13	$\delta(H_1–C_1–C_2)$	13	$\delta(H_2–C_2–C_1)$
1072	0.1	36	$\nu(C_9–C_8)$	9	$\nu(O_9=C_9)$				
984	4.4	51	$\delta(H_m–C_m–S)$	38	$\gamma(H_7–C_7=C_8–H_8)$				
978	6.1	49	$\gamma(H_7–C_7=C_8–C_1)$	35	$\delta(H_m–C_m–S)$				
973	9.8	33	$\delta(C_3–C_4–C_5)$	27	$\delta(C_6–C_2–C_1)$	12	$\nu(C_5–C_6)$		

The frequencies are shown in cm^{-1}, with their respective calculated intensities. The modes that are localized mainly on the methyl group (1457, 1435 and 1329 cm^{-1}) are excluded from the table, but if a methyl mode is strongly mixed with other modes, it is symbolized with an *m*. Only the modes contributing ~10% or more to the PED are shown.

Table 2 **Calculated frequencies (B3LYP/6–311++G(d,p), scaled by 0.9614), intensities and PED for the chromophore in the ground state, based on the 1NWZ structure.**

delocalized, as can be clearly seen from the tables. The vibrations localized mainly on the (terminating) methyl group are omitted from the tables. Table 1 and 2 list the results for the ground state configuration of the chromophore (HF in Table 1 and DFT in Table 2; both based on pdb–file 1NWZ), while Table 3 and 4 show the results of the first stable photocycle intermediate, I_0 (HF in Table 3 and DFT in Table 4; both based on pdb–file 3PYP). Before the NMA was done, the structures were geometry optimised (see the Methods section for more details on the calculation).

A comparison of the calculations presented in this chapter, and the previously obtained experimental data, is obtained by plotting the results on 1NWZ with negative intensities, and the results on 3PYP with positive ones, creating a virtual difference spectrum. The HF (Figure 2) and DFT (Figure 3) results are shown together

Freq.	Int.	PED (%)					
1703	8.0	39	$\delta(O_9{=}C_9{-}C_8)$	27	$\nu(O_9{=}C_9)$	24	$\delta(S{-}C_9{=}O_9)$
1658	2.5	100	$\delta(OH^{46}{-}O_3)$				
1634	0.9	100	$\gamma(OH^{42}{-}O_3{-}H^{46})$				
1608	11	14	$\gamma(H_2O^{42}{-}O_3)$	12	$\nu(C_1{-}C_2)$	9	$\nu(C_7{=}C_8)$
1558	4.9	27	$\nu(O_3{-}C_3)$	24	$\delta(O_3{-}C_3{-}C_2)$	23	$\delta(C_4{-}C_5{-}C_6)$
1510	22	21	$\delta(C_7{-}C_6{-}C_1)$	16	$\delta(H_7{-}C_7{=}C_8)$	12	$\delta(C_3{-}C_4{-}C_5)$
1480	13	19	$\nu(C_1{-}C_2)$	17	$\delta(H_7{-}C_7{=}C_8)$	16	$\nu(C_4{-}C_5)$
1460	1.1	41	$\nu(H_8{-}C_8)$	12	$\delta(C_8{=}C_7{-}C_6)$	11	$\delta(C_9{-}C_8{=}C_7)$
1432	0.1	24	$\delta(C_6{-}C_1{-}C_2)$	20	$\nu(H_2{-}C_2)$	11	$\delta(H_7{-}C_7{=}C_8)$
1392	0.2	17	$\nu(H_5{-}C_5)$	16	$\nu(O_3{-}C_3)$	15	$\delta(O_3{-}C_3{-}C_2)$
1321	26	43	$\delta(C_7{-}C_6{-}C_1)$	14	$\nu(C_7{=}C_8)$	10	$\delta(H_7{-}C_7{=}C_8)$/ $\delta(H_1{-}C_1{-}C_2$/ $\delta(H_4{-}C_4{-}C_5)$
1264	32	16	$\nu(H_2{-}C_2)$	14	$\delta(H_7{-}C_7{=}C_8)$	13	$\nu(H_5{-}C_5)$
1224	0.1	26	$\nu(C_2{-}C_3)$	20	$\nu(C_5{-}C_6)$	17	$\delta(C_8{=}C_7{-}C_6)$
1189	42	18	$\delta(C_3{-}C_4{-}C_5)$	12	$\nu(C_4{-}C_5)$	12	$\nu(H_8{-}C_8)$/ $\delta(H_4{-}C_4{-}C_5)$
1141	8	22	$\delta(H_1{-}C_1{-}C_2)$	12	$\nu(H_2{-}C_2)$	12	$\nu(H_5{-}C_5)$
1101	0.5	20	$\nu(H_2{-}C_2)$	13	$\delta(H_4{-}C_4{-}C_5)$	13	$\nu(H_5{-}C_5)$
1025	1.0	100	$\gamma(H_5{-}C_5{-}C_4{-}H_4)$				
1014	23	97	$\gamma(H_7{-}C_7{=}C_8{-}C_8)$				
1009	1	72	$\delta(O_9{=}C_9{-}C_8)$	15	$\nu(C_9{-}C_8)$	12	$\delta(H_m{-}C_m{-}S)$

124 The frequencies are shown in cm^{-1}, with their respective calculated intensities. The modes that are localized mainly on the methyl group (1450, 1435 and 1350 cm^{-1}) are excluded from the table, but if a methyl mode is strongly mixed with other modes, it is symbolized with an *m*. Only the modes contributing ~10% or more to the PED are shown.

Table 3 **Calculated frequencies (HF/6–311++G(p,d) scaled by 0.9054), intensities and PED for the chromophore in the I_0–state, based on the 3PYP structure.**

with the experimentally obtained I_0–pG spectrum from chapter 4 (the grey line). The NMA results demonstrate unambiguously that many possible isomerization markers are present in the experimental I_0–pG spectrum (*i.e.* many bands are characteristic for one particular isomer). The experimental difference spectrum contains, besides isomerization markers, features from changed interactions between amino acid residues in the chromophore–binding pocket, and because the used model chromophore in this work is not identical to the native protein situation, not all experimentally observed features are discussed.

1625–1750 cm^{-1}

Upon photoexcitation the chromophore isomerizes about the central double bond $C_7{=}C_8$ group, and, if it succeeds in breaking the hydrogen bond between its carbonyl group and the residue Cys69, the C=O flips to the other side and forms a biologically productive state. Upon H–bond breaking (formation/strengthening), the C=O stretch will upshift (downshift) in frequency. ν(C=O) containing

Freq.	Int.	PED (%)							
1658	2.9	89	$\delta(H_2O^{46})$						
1646	4.8	73	$\nu(O_9{=}C_9)$						
1615	1.2	79	$\delta(H_2O^{42})$	13	$\gamma(OH^{42})–O_3–H^{46})$				
1593	5.1	25	$\nu(C_1–C_2)$	14	$\nu(C_7{=}C_8)$	8	$\nu(C_4–C_5)$		
1546	20	30	$\delta(H_7–C_7{=}C_8)$	14	$\delta(C_4–C_5–C_6)$	12	$\nu(C_6–C_7)$		
1506	26	35	$\nu(O_3–C_3)$	14	$\delta(C_2–C_3–C_4)$	10	$\delta(H_7–C_7{=}C_8)$		
1475	4.7	36	$\delta(H_8–C_8{=}C_7)$	12	$\nu(C_2–C_1)$	10	$\delta(H_1–C_1–C_2)$		
1463	5.9	29	$\nu(C_4–C_5)$	20	$\delta(H_8–C_8{=}C_7)$	8	$\nu(C_5–C_6)$		
1437	13	15	$\delta(H_2–C_2–C_1)$	12	$\nu(C_2–C_3)$	12	$\delta(C_4–C_5–C_6)$	11	$\nu(C_1–C_2)$
1381	0.9	34	$\nu(O_3–C_3)$	16	$\delta(H_1–C_1–C_2)$	9	$\nu(C_6–C_7)$		
1344	0.3	18	$\delta(H_5–C_5–C_4)$	15	$\delta(H_4–C_4–C_5)$	15	$\delta(H_7–C_7{=}C_8)$		
1268	24	19	$\delta(H_1–C_1–C_2)$	18	$\delta(H_5–C_5–C_4)$	14	$\delta(H_7–C_7{=}C_8)$	14	$\delta(C_4–C_5–C_6)$
1245	4.0	24	$\nu(C_3–C_2)$	14	$\nu(C_3–C_4)$	14	$\delta(H_4–C_4–C_5)$		
1198	0.2	20	$\nu(C_6–C_7)$	18	$\delta(C_4–C_5–C_6)$	15	$\delta(H_8–C_8{=}C_7)$		
1154	1.4	22	$\delta(H_2–C_2–C_1)$	21	$\delta(H_5–C_5–C_4)$	19	$\delta(H_4–C_4–C_5)$	17	$\delta(H_1–C_1–C_2)$
1084	0.1	26	$\delta(H_2–C_2–C_1)$	14	$\delta(H_4–C_4–C_5)$	13	$\delta(H_5–C_5–C_4)$		
990	0.4	39	$\nu(C_8–C_9)$	19	$\delta(H_m–C_m–S)$	8	$\nu(O_9{=}C_9)$		
972	5.9	100	$\gamma(H_5–C_5–C_4–H_4)$						

The frequencies are shown in cm^{-1}, with their respective calculated intensities. The modes that are localized mainly on the methyl group (1423, 1413 and 1309 cm^{-1}) are excluded from the table, but if a methyl mode is strongly mixed with other modes, it is symbolized with an *m*. Only the modes contributing ~10% or more to the PED are shown.

Table 4 **Calculated frequencies (B3LYP/6–311++G(d,p), scaled by 9614), intensities and PED for the chromophore in the I_0–state, based on the 3PYP structure**.

modes are generally found in the region around 1700 cm^{-1} (42). In the experimentally determined difference spectrum in Figure 2, the features above 1720 cm^{-1} are assigned to the C=O of Glu46 (14, 43), and because we did not include the full residue (but a water molecule), these features are not observed in our calculations.

We experimentally observe a band at 1644 cm^{-1} that upshifts to 1662 cm^{-1} upon photoisomerization (grey line in Figure 2 and (5, 7)), similar to resonance Raman experiments that found an upshift from 1631 to 1666 cm^{-1} in pCA (44). The calculations show a normal mode delocalized over the $C_9{=}O$, $O_9{=}C_9–C_8$ atoms and to a minor extent over the $C_8{=}C_7–C_6$ groups at 1664 cm^{-1} for HF in the PYP ground state calculation (Table 1). From DFT computations two modes with high $\nu(C_9{=}O_9)$ stretching contribution are resolved, an intense one at 1633 cm^{-1} and a less intense mode at 1652 cm^{-1} (Table 2). A similar splitting is observed by (6), which reports bands at 1509 and 1648 cm^{-1} for the *trans* conformation and 1574 and 1696 cm^{-1} for the *cis* conformation of 3PYP (DFT; but without H–bond donor to $C_9{=}O_9$). The appearance of two distinct modes

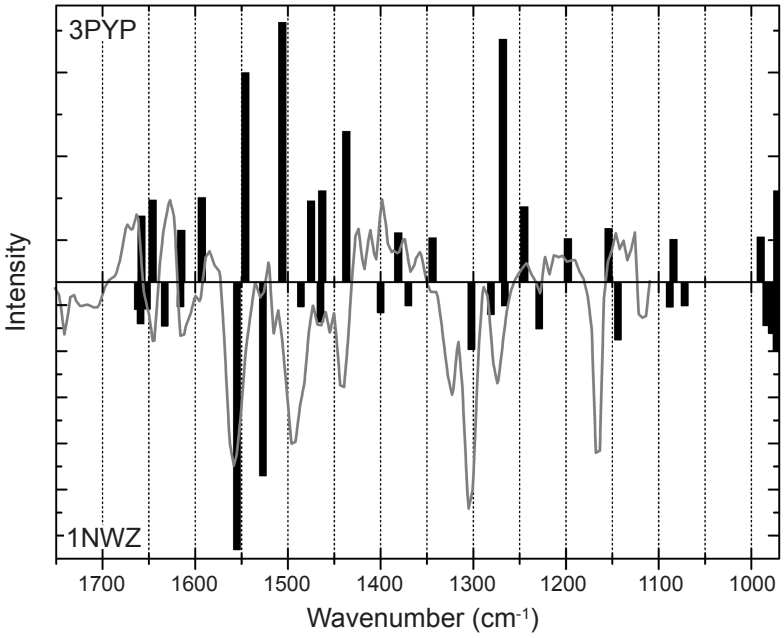


Figure 3 **Comparison of the DFT calculations on 1NWZ (bottom panel) and 3PYP (top panel), and the experimental I_0-pG difference spectrum (grey line, ΔmOD x 10 as function of wavenumber). The calculated intensities are plotted as a function of wavenumber (scaled by 0.9614), and have an offset of an additional 5 units (negative for the top panel, positive for the bottom one) to make the low intensity frequencies better visible. The shared x-axis is the baseline for the experimental spectrum.**

with high $\nu(C_9{=}O_9)$ contribution in our results is likely due to a high degree of mixing between chromophore and water vibrations. This mixing is probably overestimated by the DFT computations, and results in an underestimation of the computed $\nu(C_9{=}O_9)$ frequency compared to the experimental one. Note that the HF computations produced only one mode with a high contribution from $\nu(C_9{=}O_9)$, upshifting from 1664 to 1703 cm^{-1}, much higher in frequency than found with DFT. Calculations (29) on a different chromophore model (in the ground state; with three water molecules attached, similar to the model used here) reported $\nu(C_9{=}O_9)$ even at 1738 cm^{-1}. It is known from isotope labelled resonance Raman experiments (44) that the carbonyl stretching upshifts upon the formation of I_0, hence we can conclude that the DFT computed mode that better corresponds to the experimentally observed $\nu(C_9{=}O_9)$ is the one of the more intense mode at 1633 cm^{-1}. These modes show both in HF and DFT a clear upshift in the isomerized conformation of the chromophore (Figure 2 and 3): in HF from 1664 to 1703 cm^{-1} (Table 3) and in DFT from 1633 cm^{-1} to 1646 cm^{-1} (Table 4). The experimental band that upshifts from 1644 cm^{-1} to 1662 cm^{-1} is thus assigned to the breaking of the H–bond of the chromophore with Cys69 upon photoisomerization.

Vibrational frequencies originating from C=C stretch modes are also generally observed in the 1600 cm^{-1} region (42). Experimentally, positive features observed in this region appear at 1627 and 1583 cm^{-1}, and negative ones at 1644, 1616 and 1594 cm^{-1} (Figure 2, 3 and chapter 4). The PED reveals high contributions of the $\nu(C_7{=}C_8)$ mode at (DFT between brackets): 1664 (1616) cm^{-1} for

the ground state conformation (see Table 1 and 2) and 1608 (1593) cm^{-1} for the isomerized form (see Table 3 and 4). The 1664 cm^{-1} mode for HF as mentioned earlier is actually a (bending) mode that is mixed with modes localized around the chromophore's carbonyl C_9=O_9. The DFT results indicate a downshift of $\nu(C_7=C_8)$ upon isomerization. Accordingly, a 1633 cm^{-1} resonance Raman band (based on calculations (45, 46)) has been assigned to have a contribution of $\nu(C_7=C_8)$, and to downshift with ~10 cm^{-1} upon isomerization (30). Because of this general correspondence, and all bands with a high contribution of $\nu(C_7=C_8)$ mode are either close to bands that contain high $\nu(C_9=O_9)$ mode contribution or are mixed with them, we conclude that the experimental band at 1644 cm^{-1} is a mix of $\nu(C_7=C_8)$ and $\nu(C_9=O_9)$. Upon isomerization the $\nu(C_7=C_8)$ contribution to the 1644 cm^{-1} mode then downshifts, resulting in a contribution to the 1627 cm^{-1} band. Another report did also assign the 1644 cm^{-1} band to have a contribution from $\nu(C_7=C_8)$, but not from $\nu(C_9=O_9)$, which could possibly be caused by the use of a different chromophore model (a cysteine attached to the sulphur (6)).

In the following we will discuss the results from the DFT calculations only, because they seem to generate a better agreement with the experimental data (see also Figures 2 and 3). In addition we focus on the band positions and not on the relative intensities because the experimentally determined intensity and location can in general not be reproduced reliably (*1*).

1530–1625 cm^{-1}

From DFT calculations we find two delocalised 'backbone' modes (involving C_1–C_2, C_4–C_5, C_7=C_8, C_8–C_9 and C_9=O_9) at 1616 cm^{-1} and at 1555 cm^{-1}. These modes seem to downshift to 1593 and 1546 cm^{-1}, though also the nature of the delocalisation changes somewhat between the two configurations. Experimentally, negative features are observed at 1616, 1594 and 1555 cm^{-1}, positive features observed at 1627 and 1583 cm^{-1}. The experimental band at 1616 cm^{-1} is thus assigned to a backbone mode, downshifting to 1583 cm^{-1}. A previous report has assigned the origin of this downshifting band to be more diverse in character than just backbone (mix of C_7=C_8, C_3=O_3, C–H and amide I, (6)). The experimentally observed 1555 cm^{-1} band is also a backbone mode, but a downshifting partner band is not present. A possible cause of this absence might be the overlap of the bleach and product band in the experiment, and/or with a potential role of the lack of a high contribution of $\nu(C_9=O_9)$ to the I_0 conformation.

1100–1530 cm^{-1}

In the region 1400–1527 cm^{-1} several normal modes mainly delocalized over the phenol ring are found, whereas between

1400 and 1144 cm^{-1} similar modes include the isomerizable $C_7=C_8$ group. Upon isomerization, the chromophore changes its orientation, changing the interactions between the phenol oxygen and the hydrogen bond donating residues (5, 7, 14, 47). In the ground state, for a fairly large number of the modes a good agreement with the frequency (not intensity) of an experimentally observed band is obtained: 1555, 1527, 1486, 1465, 1370, 1302, 1281, 1266, and 1229 cm^{-1}. This suggests that the DFT approach does render a fairly good description of the ground state of the chromophore. In the I_0 configuration the DFT results are more severely off the mark, and agreement is found only for the three modes at 1245, 1198 and 1154 cm^{-1}, all phenol modes, but the 1198 cm^{-1} band has some additional contribution of $H_8–C_8–C_7$.

Possible reasons for the differences between the calculations and the experimentally observed bands were discussed above (mode mixing, overlap). Another possible source for the observed differences involving the phenol ring modes in I_0, are the structural differences between the native protein and the used model chromophore. In the protein, the two H–bonds on the phenolate oxygen are donated by two chemically different groups (Tyr42 and Glu46), which are probably forming H–bonds with the chromophore of different strength. Also, sterical constraints in the protein are certainly not reproduced in our model, where the two added water molecules are free to move during the geometry optimization. However, the ground state results should then show a similar discrepancy as the I_0 results. Possibly, differences in hydrogen bond distance play a role, which will reflect upon the vibrations modes on the inserted water molecules. A hint might be that we observe a strong 44 cm^{-1} downshift upon isomerization for the δH_2O^{42} mode, while only a small 4 cm^{-1} downshift occurs for δH_2O^{46}. The isomerization has clearly a larger effect on the interaction with H_2O^{42} than for H_2O^{46}.

To check if the changing frequencies reflect changes in H–bond distance, the final *trans* and *cis* geometries used to calculate the frequencies are studied in more detail. The structures reveal H–bond distances (in Å) for $O_3–H_2O^{42}$ of 1.80 in *trans*, and 1.96 in *cis*, and for $O_3–H_2O^{46}$ of 1.78 in *trans* and 1.73 in *cis*. The observed 44 cm^{-1} downshift for H_2O^{42} is accompanied by a H–bond distance increase (+0.16 Å), while the 4 cm^{-1} downshift for H_2O^{46} is accompanied by a decrease in H–bond distance (–0.05 Å). Obviously the waters signal the different structure of chromophore in a way that is subtler than via bond length (the same applies to the HF results). Remember also that the water molecules are free to move during the geometry optimization process, and that we know from infrared experiments on the protein that for instance the C=O stretching frequency of Glu46 changes upon the formation of I_0, reflecting changes in H–bond distance (5, 7, 14). A better understanding of the influence of the protein in

facilitating and regulating the photoisomerization process can only be accomplished if a more extended (hence more computing–intensive) model, including the nearby amino acids, is used in the calculations.

Below 1100 cm^{-1}

Although the experimental data do not extend to the region below 1100 cm^{-1}, this region requires some extra attention, as it is reported to contain a low frequency structural isomerization marker between 950–1100 cm^{-1} (6, 29, 44). Isotope labelling of C_9 with ^{13}C has shown a downshift of ~14 cm^{-1} of the 1053 cm^{-1} band, ultimately downshifting ~50 cm^{-1} upon isomerization (44), and it has been assigned to $\nu(C_8–C_9)$ (6, 29). We report that the 1072 cm^{-1} band that downshifts to 990 cm^{-1} is actually a mix of the low frequency $\nu(C_9=O_9)$ band (9 %) and $\nu(C_8–C_9)$ (36 %, see Tables 2 and 4), in accordance to what would be expected from an isotope labelling of C_9.

General remarks

A comparison of the results obtained by NMA on a model chromophore and the experimental data measured on the native protein requires that some general considerations have to be taken into account. In practice, bands have a certain spectral width and spectral overlap will occur. In addition, as can be seen in Figure 2 and 3, the calculated intensities do not match the experimentally observed infrared intensities (all bands). The use of different model chromophores (see the Introduction section), with different terminating groups at the sulphur atom end of the chromophore, or the presence/absence of hydrogen bond donors, will also influence the outcome of NMA, and highlights the sensitivity of the chromophore's normal modes to composition and interactions with its surrounding residues. The fact that our I_0 results obtained with DFT do not show a general correspondence with the positive features in the measured spectrum might have a similar origin, and may give a structural hint as to the reason behind the observed non–isomerizable nature of the chromophore with a methyl attached to the sulphur in solution (24). Additional issues might arise due to the planar conformation of the calculated chromophores, while X–ray diffraction shows it to be strained. The real–time probed structure in the picosecond IR experiments might therefore differ from the cryotrapped X–ray structure. Structural optimization of the 3PYP coordinates can also have resulted in a too planar structure that differs significantly from the strained (since further relaxation to I_1 takes place on the 0.7–1 ns time scale) structure of the I_0 intermediate. The planarity influences the degree of π–conjugation and has an effect on the frequencies of both the phenyl ring and of the carbon chain moieties. X–ray diffraction has revealed that the chromophore is twisted in the

ground state (the carbonyl has an angle of ~23° with the plane formed by the chromophore's ring and central C_7=C_8 bond, (19)) and even more in the early cryotrapped intermediate (now 38° for 3PYP (16)). *In vivo* molecules thermally vibrate however, and a direct comparison between the results from calculations (that make it planar due to the lack of interactions), X–ray diffraction data (measured on frozen crystals), and native protein is therefore not straightforward. Alternatively, differences might arise from the use of a different basis set and/or computational method. A further source of disparity between experiment and theory is partly caused by anharmonicity of the experimental frequencies (48) and differences in chromophore orientation between the chromophore within the protein and in vacuum, and of course by the omission of the interactions between chromophore and protein. All of the mentioned factors can result in significant frequency shifts.

A decrease of the existing differences between vibrations that are observed in theory and experiment can be obtained by using multiple scaling factors, as some basis sets structurally over– or underestimate certain bond lengths (49). In any case, NMA in combination with PED is a helpful tool in deciding where to start with various experimental approaches.

Conclusions

For all calculations presented here it is clear that practically all chromophore vibrational modes are strongly coupled. The delocalisation of vibrational modes renders the assignment of the experimentally observed frequencies difficult. However, this should not prohibit a correct assignment of isomerization markers. The use of *ab initio* and hybrid DFT calculations in combination with PED greatly enhances the understanding of experimental data and the objective assignment of bands. The agreement between experimental data and calculations is closer for DFT than for HF. We find that DFT on the ground state structure accurately predicts the vibrational properties of the chromophore. However, the DFT calculations on I_0 seem to correspond much less favourably. Either the 3PYP structure representing I_0 does not match to the transiently probed chromophore configuration, or the calculation method has not adequately predicted the frequencies. Ultimately, all residues that interact with the chromophore should be included for an accurate calculation of the normal modes in the infrared fingerprint region.

Acknowledgements

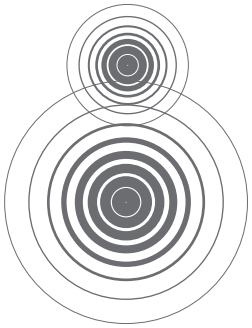
E. Romero is gratefully acknowledged for allowing use of her desktop pc, and Zbigniew Kisiel for helpful discussions and for supplying and updating the VIBCA program.

References

1. Cramer, C. J. (2002) *Essentials of computational chemistry* (John Wiley & Sons Ltd., Chichester, England).
2. Hellingwerf, K. J., Hendriks, J. & Gensch, T. (2003) *Journal of Physical Chemistry A* 107, 1082–1094.
3. Cusanovich, M. A. & Meyer, T. E. (2003) *Biochemistry* 42, 4759–4770.
4. Larsen, D. S. & van Grondelle, R. (2005) *Chemphyschem* 6, 828–837.
5. Groot, M.–L., van Wilderen, L. J. G. W., Larsen, D. S., van der Horst, M. A., van Stokkum, I. H. M., Hellingwerf, K. J. & van Grondelle, R. (2003) *Biochemistry* 42, 10054–10059.
6. Heyne, K., Mohammed, O. F., Usman, A., Dreyer, J., Nibbering, E. T. J. & Cusanovich, M. A. (2005) *Journal of the American Chemical Society* 127, 18100–18106.
7. van Wilderen, L. J. G. W., van der Horst, M. A., van Stokkum, I. H. M., Hellingwerf, K. J., van Grondelle, R. & Groot, M.–L. (2006) *Proc Natl Acad Sci U S A* 103, 15050–15055.
8. Groenhof, G., Lensink, M. F., Berendsen, H. J. C., Snijders, J. G. & Mark, A. E. (2002) *PROTEINS: Structure, Function, and Genetics* 48, 202–211.
9. Groenhof, G., Lensink, M. F., Berendsen, H. J. C. & Mark, A. E. (2002) *PROTEINS: Structure, Function, and Genetics* 48.
10. Larsen, D. S., van Stokkum, I. H. M., Vengris, M., van der Horst, M. A., de Weerd, F. L., Hellingwerf, K. J. & van Grondelle, R. (2004) *Biophysical Journal* 87, 1858–1872.
11. Chosrowjan, H., Mataga, N., Nakashima, N., Yasushi, I. & Tokunaga, F. (1997) *Chemical Physics Letters* 270, 267–272.
12. Mataga, N., Chosrowjan, H., Shibata, Y., Imamoto, Y. & Tokunaga, F. (2000) *Journal of Physical Chemistry B* 104, 5191–5199.
13. Brudler, R., Rammelsberg, R., Woo, T. T., Getzoff, E. D. & Gerwert, K. (2001) *Nat Struct Biol* 8, 265–70.
14. Xie, A., Hoff, W. D., Kroon, A. R. & Hellingwerf, K. J. (1996) *Biochemistry* 35, 14671–14678.
15. Imamoto, Y., Shirahige, Y., Tokunaga, F., Kinoshita, T., Yoshihara, K. & Kataoka, M. (2001) *Biochemistry* 40, 8997–9004.
16. Genick, U. K., Soltis, S. M., Kuhn, P., Canestrelli, I. L. & Getzoff, E. D. (1998) *Nature* 392, 206–9.
17. Perman, B., Srajer, V., Ren, Z., Teng, T., Pradervand, C., Ursby, T., Bourgeois, D., Schotte, F., Wulff, M., Kort, R., Hellingwerf, K. & Moffat, K. (1998) *Science* 279, 1946–50.
18. Ihee, H., Rajagopal, S., Srajer, V., Pahl, R., Anderson, S., Schmidt, M., Schotte, F., Anfinrud, P. A., Wulff, M. & Moffat, K. (2005) *Proceedings of the National Academy of Sciences of the United States of America* 102, 7145–7150.

19. Borgstahl, G. E., Williams, D. R. & Getzoff, E. D. (1995) *Biochemistry* 34, 6278–87.
20. Dux, P., Rubinstenn, G., Vuister, G. W., Boelens, R., Mulder, F. A., Hard, K., Hoff, W. D., Kroon, A. R., Crielaard, W., Hellingwerf, K. J. & Kaptein, R. (1998) *Biochemistry* 37, 12689–99.
21. Premvardhan, L. L., van der Horst, M. A., Hellingwerf, K. J. & van Grondelle, R. (2003) *Biophysical Journal* 84, 3226–3239.
22. Ujj, L., Devanathan, S., Meyer, T. E., Cusanovich, M. A., Tollin, G. & Atkinson, G. H. (1998) *Biophysical Journal* 75, 406–12.
23. Changenet–Barret, P., Espagne, A., Charier, S., Baudin, J. B., Jullien, L., Plaza, P., Hellingwerf, K. J. & Martin, M. M. (2004) *Photochemical & Photobiological Sciences* 3, 823–829.
24. Larsen, D. S., Vengris, M., van Stokkum, I. H. M., van der Horst, M. A., de Weerd, F. L., Hellingwerf, K. J. & van Grondelle, R. (2004) *Biophysical Journal* 86, 2538–2550.
25. Lee, I. R., Lee, W. & Zewail, A. H. (2006) *Proceedings of the National Academy of Sciences of the United States of America* 103, 258–262.
26. Ryan, W., Gordon, D. J. & Levy, D. H. (2002) *Journal of American Chemical Society* 124, 6194–6201.
27. Imamoto, Y., Koshimizu, H., Mihara, K., Hisatomi, O., Mizukami, T., Tsujimoto, K., Kataoka, M. & Tokunaga, F. (2001) *Biochemistry* 40, 4679–85.
28. Yoo, H. Y., Boatz, J. A., Helms, V., McCammon, J. A. & Langhoff, P. W. (2001) *Journal of Physical Chemistry B* 105, 2850–2857.
29. Usman, A., Mohammed, O. F., Heyne, K., Dreyer, J. & Nibbering, E. T. J. (2005) *Chemical Physics Letters* 401, 157–163.
30. Unno, M., Kumauchi, M., Sasaki, J., Tokunaga, F. & Yamauchi, S. (2000) *Journal of the American Chemical Society* 122, 4233–4234.
31. Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S. J., Windus, T. L., Dupuis, M. & Montgomery, J. A. (1993) *Journal of Computational Chemistry* 14, 1347–1363.
32. Berman, H. M., Henrick, K. & Nakamura, H. (2003) *Nature Structural Biology* 10, 980.
33. Scott, A. P. & Radom, L. (1996) *Journal of Physical Chemistry* 100, 16502–16513.
34. Ma, B., Schaefer, H. F. I. & Allinger, N. L. (1998) *Journal of American Chemical Society* 120, 3411.
35. Lii, J.–H., Ma, B. & Allinger, N. L. (1999) *Journal of Computational Chemistry* 20, 1593.

36. Miller, M. D., Jensen, F., Chapman, O. L. & Houk, K. N. (1989) *Journal of Physical Chemistry* 93, 4495–4502.
37. Salam, A., Deleuze, M. S. & Francois, J. P. (2001) *Chemical Physics* 271, 17–30.
38. Kisiel, Z. (2007) in *Programs for ROtational SPEctroscopy* (http://info.ifpan.edu.pl/~kisiel/prospe.htm.
39. Bode, B. M. & Gordon, M. S. (1998) *J. Mol. Graphics Mod.* 16, 133–138.
40. Morino, Y. & Kuchitsu, K. (1952) *Journal of Chemical Physics* 20, 1809.
41. Aroca Munos, R., Panchenko, Y. N., Koptev, G. S. & Stepanov, N. F. (1970) *Journal of Applied Spectroscopy* 12, 428–429.
42. Barth, B. (2000) *Progress in Biophysics & Molecular Biology* 74, 141–173.
43. Imamoto, Y., Mihara, K., Hisatomi, O., Kataoka, M., Tokunaga, F., Bojkova, N. & Yoshihara, K. (1997) *J Biol Chem* 272, 12905–8.
44. Unno, M., Kumauchi, M., Sasaki, J., Tokunaga, F. & Yamauchi, S. (2002) *Biochemistry* 41, 5668–5674.
45. Zhou, Y., Ujj, L., Meyer, T. E., Cussanovich, M. A. & Atkinson, G. H. (2001) *Journal of Physical Chemistry A* 105, 5719–5726.
46. Kim, M., Mathies, R. A., Hoff, W. D. & Hellingwerf, K. J. (1995) *Biochemistry* 34, 12669–72.
47. Devanathan, S., Brudler, R., Hessling, B., Woo, T. T., Gerwert, K., Getzoff, E. D., Cusanovich, M. A. & Tollin, G. (1999) *Biochemistry* 38, 13766–72.
48. Hout, R. F., Levi, B. A. & Hehre, W. J. (1982) *Journal of Computational Chemistry* 3, 234–250.
49. Grunenberg, J. & Herges, R. (1997) *Journal of Computational Chemistry* 18, 2050–2059.



Summary, List of Publications and Nawoord

08

Abstract

This chapter serves to provide a summary to the research presented in this thesis. In addition, a list of publications and some final words are given to conclude this thesis.

Summary

To understand in detail how proteins can generate interpretable signals about the physical conditions in our environment is a formidable challenge. In the natural sciences, photoactive proteins have generally been accepted as the most suitable model systems for this challenge. Photoactive proteins contain pigments (colour–absorbing molecules or chromophores) embedded in a protein matrix. The chromophores absorb light, and translate it by means of a certain mechanism into biological activity. The main experimental tool used to study this mechanism is picosecond visible pump–mid–infrared probe spectroscopy. This technique allows for the detection of structural changes in molecules, such as isomerization of a chromophore, or changes in hydrogen bond network interactions. For example, the change in the vibrational frequency of one single atom bond due to a changed interaction with its environment can be detected. It is the perfect tool to investigate biological systems such as proteins, and in this work it was used to study fundamental biological processes such as proton transfer, isomerization, hydrogen bond dynamics and ligand dissociation. In this work, three different proteins are studied on an ultrafast timescale (from picoseconds to nanoseconds): Green Fluorescent Protein (GFP), Photoactive Yellow Protein (PYP), and FixL. In addition, quantum chemical calculations on the chromophore of PYP are presented to obtain a better understanding of the occurring photo–induced processes.

In chapter 2 the conducted ultrafast time–resolved experiments on GFP from the jellyfish *Aequorea vicotoria* are described. Our results show unambiguously that the proton transfer process in GFP does *not* occur according to the generally accepted model. It is widely known that the colour of the fluorescent light emitted by GFP depends on a light–induced proton transfer reaction that occurs in a 'proton–wire', formed by the chromophore, a water molecule (W22), S205 and E222. Freshly synthesized GFP in the ground state contains a protonated chromophore (*i.e.* neutral), absorbs in the near UV and emits blue fluorescence, whereas upon excited state (ES) formation the emitted fluorescence shifts to the green part of the spectrum.

We have studied GFP's ES dynamics using ultrafast visible/mid–infrared pump–probe spectroscopy to elucidate the dynamics of proton transfer in the proton–wire. Model calculations based on available crystal structures proposed pathways starting from the chromophore or starting from the end of the wire, *i.e.* from the deprotonated E222. Our data show that, following optical excitation, a rapid (3 ps) protonation of E222 occurs, prior to chromophore deprotonation. The chain of proton transfer reactions starts therefore at the acceptor *end* of the wire. The anionic chromophore appears bi–exponentially with time constants of 8 and 165 ps. Our

conclusions corroborate recent electronic structure calculations done by other groups, who found that 'pulling' a proton over the wire, initiated by proton transfer from the serine to the glutamate, is energetically most favourable. Structures similar to GFP's proton–wire may have a wider role in Nature than only in (green) fluorescent proteins, as successful pathways are generally conserved.

The largest part of this thesis is focussed on PYP (chapters 3, 4, 6 and 7). Picosecond infrared spectroscopy is used to study the native wild type protein and the E46Q PYP mutant (chapters 3 and 4). PYP is a protein found in the swimming bacterium *Halorhodospira halophila*. The bacterium requires light to perform photosynthesis and to survive, but too much (blue) light is harmful. The bacterium is negatively phototactic, which means that it is able to sense the intensity of light and, according to its needs, swim towards or away from it. PYP is thought to be the mediator between the environment and an interpretable signal for the bacterium. It follows a photocycle with distinct intermediate steps, each with a different conformation and its own spectroscopic properties. The negatively charged chromophore is embedded in a hydrogen bonding network with Glu46, Tyr42 (both donating a H–bond to the phenol) and Cys69 (donating a H–bond to the chromophore's C=O).

The work performed in this thesis has led to a nearly complete picture of the ultrafast start of the photocycle, from photoexcitation up to 3 ns. Upon (blue) light excitation, the negative charge on the phenolic ring of the chromophore moves towards the carbonyl. The charge translocation rearranges the π–conjugation, rendering the C=O of the chromophore more single bond (*i.e.* C–O; chapter 3, 4 and 6). The rearrangement may facilitate the picosecond *trans* to *cis* isomerization (characterized by the appearance and disappearance of distinct normal modes; chapter 7). At the same time, the rearrangement also increases the hydrogen bond strength between the chromophore's C=O (now more C–O) and Cys69, possibly hampering the isomerization. The charge translocation and the isomerization process are also reflected upon changes in hydrogen bond interaction between the phenol and Glu46, but the hydrogen bond remains intact throughout the probed time scale in our experiments (3 ns; chapter 3 and 4). Once the *p*–coumaric acid chromophore is isomerized, *and* the hydrogen bond between the carbonyl of the chromophore and the protein backbone (*i.e.* Cys69) is lost, the biologically productive I_0 state will be formed (chapter 4). The formation of I_1 from I_0 is characterized by a 90–100% yield and we therefore conclude that I_0 is a stable long–living *cis* ground state configuration, which structurally relaxes on the nanosecond time scale to form I_1 (chapter 3). However, if the hydrogen bond between Cys69 and the chromophore's C=O

fails to break, a *cis* ground state intermediate is formed in a few picoseconds (chapter 4). The breakage of the hydrogen bond is therefore an independent process of the isomerization mechanism. A biologically productive state is only formed in about 30% of the photocycle attempts, while about 60% forms the ground state intermediate (the remaining 10% thermally decays into the ground state from ES). The failed photocycle attempts then reform the ground state, and are ready to give it another try.

The combination of charge translocation and isomerization upon light absorption seems to be a common theme in photosensors. The protein appears to play an active role in combining the two, directing the photocycle, since it not only stabilizes the negative charge on the phenolic oxygen in the ground state by an extensive hydrogen bond network, but also the chromophore's conformation (by restraining the chromophore's carbonyl). A detailed understanding of the chemical kinetic mechanism of a protein, along with a chemical basis for its efficiency, is important for the understanding of protein function, and leads to a better understanding of human metabolism.

In chapter 5, the heme–binding protein FixL is studied. We have focussed on *Bj*FixL, the oxygen sensor from the bacterium *Bradyrhizobium japonicum*. The heme has a central iron atom, where a diatomic molecule (the ligand) can bind, and is therefore similar to myoglobin and hemoglobin, the oxygen carriers for the cells in the human body. *Bj*FixL consists of two separate domains, a heme–PAS domain and a signalling kinase domain. Upon binding of oxygen, conformational changes occur in the heme–binding domain, which in turn forms a biological signal for the kinase domain. The ligand photodissociation properties of the heme binding protein *bj*FixLH are studied with picosecond visible pump–mid–infrared probe spectroscopy. The diatom CO is used as a ligand, which upon photoexcitation with a visible laser pulse is released from the heme. The infrared frequency of the CO molecule is very sensitive to interactions with the surrounding protein, and acts therefore as a probe for the chromophore–binding pocket dynamics. The infrared difference spectra indicate that the escape of photolyzed CO to solvent is preceded by transient docking within the protein in a manner similar to globins. Two orientations of the CO appear to be involved in this docking, influenced by the energy of the pump laser pulse. On a picosecond timescale, protein rearrangements may have decreased the rotational constrainedness of the CO molecule. On a nanosecond timescale, further protein relaxation may have caused the CO molecule to migrate to a new site, changed its rotational orientation, or changed its pocket conformation. Due to its similarity with the globins, but having a completely different protein (PAS–)fold, this mechanism might therefore comprise of a general motif in heme–binding

proteins.

The results discussed in this thesis have increased our understanding of GFP, PYP and FixL. These proteins are used by Nature to perceive a signal from a changing environment, and ultimately generate an interpretable signal to the cell or bacterium. Signal transduction in the studied proteins is initiated by ultrafast mechanisms common in Nature, such as proton transfer, isomerization, changes in hydrogen bonding, and ligand dissociation. The understanding of protein function is important as it is related to health and disease, and could potentially lead to improved medicine.

Samenvatting

Ultrasnelle eiwitdynamica onthuld door infraroodspectroscopie

Het is een geweldige uitdaging om in detail te begrijpen hoe eiwitten een interpreteerbaar signaal afgeven over veranderende fysieke omstandigheden. Deze uitdaging vereist modelsystemen, en in de natuurwetenschappen worden lichtactieve eiwitten over het algemeen beschouwd als de meest geschikte modellen. Lichtactieve eiwitten bevatten pigmenten (kleurabsorberende moleculen of chromoforen) die omsloten worden door een eiwitmatrix. De chromofoor absorbeert licht, wat via een bepaald mechanisme wordt omgezet in biologische activiteit. De meest gebruikte techniek in dit proefschrift is picoseconde zichtbare pomp–mid–infrarood probe spectroscopie. Met deze techniek kunnen structurele veranderingen in moleculen worden waargenomen, zoals de isomerisatie van een chromofoor, of veranderende interacties in een waterstofbrug netwerk. Bijvoorbeeld, een verandering van de vibrationele frequentie van een enkele atoombinding kan worden gedetecteerd als gevolg van het verbreken van een waterstofbrug. Het is de ideale techniek om biologische systemen zoals eiwitten te bestuderen, en wij hebben dit gebruikt om fundamentele biologische processen te onderzoeken die uiteindelijk leiden tot biologische activiteit. Deze processen bestaan uit protonenoverdracht, isomerisatie, waterstofbrug dynamica en ligand dissociatie. In het werk dat voor u ligt zijn drie verschillende eiwitten bestudeerd op een ultrasnelle tijdschaal (van picoseconden tot nanoseconden), namelijk Green Fluorescent Protein (GFP), Photoactive Yellow Protein (PYP), en FixL. Tevens worden kwantum chemische berekeningen van de chromofoor van PYP gepresenteerd, om een beter begrip te verkrijgen van de optredende lichtgeïnduceerde processen.

In hoofdstuk 2 worden de ultrasnelle tijdsopgeloste experimenten van GFP, afkomstig van de kwalsoort *Aequorea victoria*, beschreven. Onze resultaten laten onweerlegbaar zien dat het proces van protonenoverdracht in GFP *niet* het algemeen erkende model volgt. Het is algemeen bekend dat de kleur van het fluorescerende licht dat GFP uitzendt veroorzaakt wordt door een lichtgeïnduceerde protonenoverdrachtsreactie in een ‘protonendraad’, gevormd door de chromofoor, een watermolecuul (W22), een serine (S205) en een glutamaat (E222). Vers gesynthetiseerd GFP in de grondtoestand bevat een geprotoneerde chromofoor (dus elektrisch neutraal), het absorbeert in het nabij UV, terwijl bij de vorming van de aangeslagen toestand (ES) de fluorescentie naar het groene deel van het spectrum schuift.

Wij hebben de dynamica van GFP’s–ES bestudeerd door middel van picoseconde zichtbare pomp–mid–infrarood probe spectroscopie, teneinde de dynamica van de protonenoverdracht in de protonendraad op te helderen. Modelberekeningen, gebaseerd

op beschikbare kristalstructuren, hebben routes voorgesteld die ofwel starten bij de chromofoor, ofwel bij het einde van de draad, namelijk bij het gedeprotoneerde E222. Onze data laten zien dat er, als gevolg van het optisch aanslaan, een snelle (3 ps) protonering optreedt van E222, nog voordat de chromofoor deprotoneerd. De keten van protonenoverdrachtsreacties start dus bij het *einde* van de draad. De anionische chromofoor verschijnt bi–exponentieel met tijdsconstanten van 8 en 165 ps. Onze conclusies worden gestaafd door recente elektronische structuur berekeningen van andere groepen, die hebben laten zien dat het 'trekken' van een proton over de draad, geïnitieerd door een protonenoverdracht van S205 naar E222, energetisch het meest gunstig is. Structuren, lijkend op de protonendraad van GFP, spelen mogelijk een grotere rol in de natuur dan alleen maar bij (groen) fluorescerende eiwitten, omdat succesvolle routes in het algemeen worden behouden.

Het grootste gedeelte van dit proefschrift is gebaseerd op PYP (hoofdstukken 3, 4, 6 en 7). Picoseconde infrarood spectroscopie wordt gebruikt om het natuurlijk voorkomende eiwit en de E46Q PYP mutant te bestuderen (hoofdstuk 3 en 4). PYP is een eiwit dat aangetroffen wordt in de zwemmende bacterie *Halorhodospira halophila*. De bacterie heeft licht nodig om fotosynthese te kunnen doen en om te overleven, maar teveel (blauw) licht is schadelijk. De bacterie vertoond negatieve phototaxis, wat betekend dat het in staat is om de intensiteit van licht waar te nemen en daar juist naar toe of vanaf te zwemmen, al naar gelang zijn behoefte. PYP wordt verondersteld de mediator te zijn tussen de omgeving en een interpreteerbaar signaal voor de bacterie. Het volgt een fotocyclus met afzonderlijke, tussenliggende stappen, elk met een eigen conformatie en spectrale eigenschappen. De negatief geladen chromofoor ligt ingebed in een waterstofbrug netwerk van Glu46, Ty42 (beiden waterstofbrugdonoren aan de fenol) en Cys69 (waterstofbrugdonor aan de C=O van de chromofoor).

Het uitgevoerde werk in dit proefschrift heeft geleid tot een nagenoeg compleet beeld van de ultrasnelle start van de fotocyclus, vanaf lichtexcitatie tot 3 ns. Als gevolg van (blauw) licht excitatie verplaatst de negatieve lading op de fenolring van de chromofoor zich in de richting van de carbonyl. De ladingsverplaatsing herschikt de π–conjugatie, en geeft de C=O van de chromofoor het karakter van een enkelvoudige binding (hoofdstuk 3, 4 en 6). De herschikking zou de *trans* naar *cis* isomerisatie (gekenmerkt door het verschijnen en verdwijnen van karakteristieke eigentrillingen, zie hoofdstuk 7), die op een picoseconde tijdschaal plaatsvindt, kunnen vergemakkelijken. Tegelijkertijd wordt de waterstofbrug tussen de C=O van de chromofoor (nu meer C–O) en Cys69 versterkt, wat de isomerisatie mogelijk tegenwerkt. De ladingsverschuiving en het isomerizatieproces heeft ook zijn weerslag op veranderingen in waterstofbrug interacties tussen de fenol en Glu46, maar deze

waterstofbrug blijft intact gedurende de gehele onderzochte tijdschaal in onze experimenten (3 ns; hoofdstuk 3 en 4). Zodra *zowel* de para–hydroxy–kaneelzuur chromofoor is geïsomeriseerd, *als* de waterstofbrug tussen de carbonyl van de chromofoor en de eiwit ruggengraat (oftewel Cys69) is verbroken, dan pas wordt de biologisch productieve I_0 toestand gevormd (hoofdstuk 4). De vorming van I_1 uit I_0 wordt gekenmerkt door een 90–100% opbrengst. Vanwege deze opbrengst concluderen wij dat I_0 een stabiele langlevende grondtoestandsconfiguratie heeft, die in structuur relaxeert op een nanoseconde tijdschaal om I_1 te vormen (hoofdstuk 3). Echter, als de waterstofbrug tussen Cys69 en de C=O van de chromofoor niet verbroken wordt, wordt er een *cis* grondtoestands–tussenvorm gevormd in een paar picoseconden (hoofdstuk 4). De verbreking van de waterstofbrug is dus een proces dat onafhankelijk is van het isomerisatie mechanisme. Een biologisch productieve toestand wordt alleen gevormd in slechts ongeveer 30% van de fotocycluspogingen, terwijl ongeveer 60% de grondtoestands–tussenvorm vormt (de rest valt direct terug naar de grondtoestand vanuit ES). De mislukte fotocyclus–pogingen vormen dan opnieuw de grondtoestand, en kunnen het nog een keer proberen.

De combinatie van ladingsverschuiving en isomerisatie als gevolg van lichtabsorptie lijkt een veelvoorkomend thema te zijn bij fotosensoren. Het eiwit blijkt een actieve rol te spelen door het combineren van die twee, en dus stuurt het de fotocyclus aan; het stabiliseert de negatieve lading op de fenolzuurstof in de grondtoestand door een uitgebreid waterstofbrug netwerk, en het stabiliseert de chromofoor conformatie in de grondtoestand door de carbonyl van de chromofoor vast te houden. Een gedetailleerd begrip van de chemische kinetiek van een eiwit (samen met een chemische basis voor haar efficientie) is belangrijk voor het begrijpen van eiwit functie in het algemeen, en leidt tot een beter begrip van de stofwisseling in de mens.

In hoofdstuk 5 wordt het heem–groep bindende eiwit FixL bestudeerd. Wij hebben *Bj*FixL onderzocht, de zuurstof sensor van de *Bradyrhizobium japonicum* bacterie. De heem heeft een centraal gelegen ijzeratoom waar een diatomisch molecuul (de ligand) kan worden vastgebonden, en lijkt dus op myoglobine en hemoglobine, de zuurstofdragers voor de cellen van het menselijk lichaam. *Bj*FixL bestaat uit twee afzonderlijke domeinen; een heem–PAS domein, en een signaalafgevend kinase domein. Bij het binden van zuurstof vinden er vormveranderingen plaats in het heem–domein, welke op hun beurt een biologisch signaal vormen voor het kinase domein. De lichtdissociatie eigenschappen van het heem–bindende eiwit *Bj*FixLH worden bestudeerd met picoseconde zichtbare pomp–mid–infrarode probe spectroscopie. Het diatoom CO wordt gebruikt als ligand, en zijn verbinding

met de heem wordt verbroken als gevolg van lichtexcitatie. De infrarode frequentie van het CO molecuul is erg gevoelig voor interacties met het omliggende eiwit, en funktioneerd daardoor als een sonde voor de dynamica van de chromofoor–bindende eiwitholte. De infrarood verschilspectra wijzen erop dat het ontsnappen van het losgeschoten CO naar de oplossing vooraf gegaan wordt door het tijdelijk dokken van de CO binnen het eiwit, op een manier die vergelijkbaar is met de globines. Er lijken twee oriëntaties van de CO deel te nemen in dit dokkingsproces. Op een picoseconde tijdschaal vindt er een herschikking van het eiwit plaats, wat waarschijnlijk rotationele vrijheid van het CO molecuul heeft beperkt. In een nanoseconde vindt een verdere relaxatie van het eiwit plaats, wat ervoor zorgt dat oftewel het CO molecuul naar een nieuwe dokplaats migreerd, de rotationele orientatie heeft veranderd, of de configuratie van de dokplaats. Door de overeenkomsten met de globines enerzijds, maar met een compleet andere opgevouwen structuur anderzijds, zou het zo kunnen zijn dat het beschreven mechanisme een wijdverbreid motief is in heem–gebonden–eiwitten.

De resultaten bediscussieerd in dit proefschrift hebben onze kennis van GFP, PYP en FixL vergroot. Deze eiwitten worden door de natuur gebruikt om signalen waar te nemen van een veranderende leefomgeving, en deze uiteindelijk om te zetten in interpreteerbare signalen voor de cel of bacterie. Signaal transductie in de bestudeerde eiwitten wordt geïnitieerd door algemeen in de natuur voorkomende ultrasnelle mechanismen, zoals protonenoverdracht, isomerisatie, veranderingen in waterstofbruggen en ligand dissociatie. Begrijpen hoe eiwitten functioneren is belangrijk, omdat het gerelateerd is aan gezondheid en ziekte, en het potentie implicaties heeft voor de ontwikkeling van verbeterde medicijnen.

List of Publications

1. Dynamics of carbon monoxide photodissociation in *Bradyrhizobium Japonicum* FixL probed by picosecond mid-infrared spectroscopy
van Wilderen LJGW, Key J, Van Stokkum IHM, Moffat K, van Grondelle R, Groot ML *In preparation*

2. (Ultra)Fast spectroscopy of Xanthopsins
van Wilderen LJGW, Groot ML in Photoreceptor proteins from purple bacteria (N. Hunter, ed.) *In preparation*

3. The pathway for proton transfer in Green Fluorescent Protein
van Wilderen LJGW, Van Stokkum IHM, Hellingwerf KJ, van Grondelle R, Groot ML *Submitted*

4. Femtosecond dispersed and time-resolved infrared spectroscopy on proteins
Groot ML, **van Wilderen LJGW**, Didonato M Photochemical and Photobiological Sciences Special Issue *to be published*

5. Ultrafast infrared Spectroscopy reveals key step for successful entry into photocycle for Photoactive Yellow Protein
van Wilderen LJGW, van der Horst MA, Van Stokkum IHM, Hellingwerf KJ, van Grondelle R, Groot ML Proceedings of the National Academy of Sciences of the USA 2006 103 (41) 15050-15055

6. Initial electron donor and acceptor in isolated Photosystem II reaction centers identified with femtosecond mid-IR spectroscopy
Groot ML, Pawlowicz NP, **van Wilderen LJGW**, Breton J, van Stokkum IHM, van Grondelle R Proceedings of the National Academy of Sciences of the USA 102 (37) 13087–13092 2005

7. Initial charge separation in photosystem II reaction centers identified with femtosecond mid-infrared spectroscopy
Groot ML, Pawlowicz NP, **van Wilderen LJGW**, Breton J, van Stokkum IHM, van Grondelle R
FEBS Journal 272 449-450 Suppl. 1, JUL 2005

8. Early steps in blue light reception by plants: an ultrafast mid-infrared spectroscopic study of the LOV2 domain of Phototropin
Alexandre MTA, **van Wilderen LJGW**, van Grondelle R, Hellingwerf KJ, Groot ML, Kennis JTM Biophysical Journal 88 (1): 509A-509A Part 2 Suppl. S, JAN 2005

9. Femtosecond visible/mid-IR pump-probe study of the Photosystem II core antenna complex CP47
Groot ML, Breton J, **van Wilderen LJGW**, Dekker JP, van Gron-

delle R Journal of Physical Chemistry B 108 8001-8006 2004

10. Femtosecond visible pump – midinfrared probe setup for the study of protein dynamics. Isomerization in Photoactive Yellow Protein
Groot ML, **van Wilderen LJGW**, Larsen DS, van der Horst MA, van Stokkum IHM, Hellingwerf KJ, van Grondelle R in *Femtochemistry and Femtobiology: Ultrafast Events in Molecular Science* (M.M. Martin and J.T. Hynes, eds.) Elsevier, Amsterdam 417-420 2004

11. Initial steps of signal generation in Photoactive Yellow Protein revealed with femtosecond mid-infrared spectroscopy
Groot ML, **van Wilderen LJGW**, Larsen DS, van der Horst MA, van Stokkum IHM, Hellingwerf KJ, van Grondelle R Biochemistry 42 (34) 10054-10059 2003

12. In-situ study of pearlite nucleation and growth during isothermal austenite decomposition in nearly eutectoid steel
Offerman SE, **van Wilderen LJGW**, van Dijk NH, Sietsma J, Rekveldt MT, van der Zwaag S Acta Materialia 51 (13) 3927-3938 2003

13. Cluster formation of pearlite colonies during the austenite/pearlite phase transformation in eutectoid steel
Offerman SE, **van Wilderen LJGW**, van Dijk NH, Sietsma J, Rekveldt MT, van der Zwaag S Physica B 335 (1-4) 99-103 2003

14. Neutron depolarization study of the austenite/pearlite phase transformation in steel
van Wilderen LJGW, Offerman SE, van Dijk NH, Sietsma J, Rekveldt MT,van der Zwaag S Applied Physics A-Materials Science and Processing 74 S1052-S1054 Part 2 2002

15. Copper doping of silicate glasses by the ion-exchange technique: A photoluminescence spectroscopy study
Borsella E, Dal Vecchio A, Garcia MA, Sada C, Gonella F, Polloni R, Quaranta A, **van Wilderen LJGW** Journal of Applied Physics 91 (1) 90-98 2002

Nawoord

Ce l'ho fatta!!

Na vier en een half jaar lezen, meten, analyseren en schrijven ligt het boekje er toch! Het is een leerzame en intensieve periode geweest. Een aantal projecten zijn gestart en goed geëindigd, maar andere ook gestaakt… Maar, that's all in the game! Graag wil ik mijn proefschrift besluiten met het bedanken van al diegenen die mij hebben bijgestaan of geholpen gedurende deze periode.

Ten eerste Rienk, die het aan heeft gedurfd om een materiaakundig ingenieur los te laten op een verzameling nietsvermoedende eiwitten. Bedankt voor je steun, inspiratie, en voor het benadrukken van het belang van openheid in de wetenschap!

Dan wil ik uiteraard ook Professor Marloes bedanken, die mij van veel dichterbij heeft gevolgd. Jij hebt mij begeleid bij mijn eerste stappen in de wereld van eiwitten en infrarood. Van jou heb ik eigenlijk alles geleerd om een zelfstandig wetenschapper te worden: van experimenteren, analyseren en interpreteren, tot het gecondenseerd en begrijpelijk opschrijven ervan. Jouw wetenschappelijke aanpak, kennis van zaken, en jouw positieve instelling hebben het tot een plezier gemaakt om met jou samen te werken! Bedankt voor alles!

Verder wil ik ook graag Ivo bedanken, wiens software voor mij zelfs na vier jaar ermee gestoeid te hebben nog steeds geheimen heeft! Ongelofelijk wat jij met onze infrarood data hebt weten toe doen. Zonder jouw hulp hadden we nooit zulke mooie resultaten behaald. Ook wil ik Jos, Henny en Sandrine bedanken, die altijd klaar staan voor iedereen in het lab, en tenslotte ook Petra, Yvonne en Claudia voor hun administratieve steun.

Tevens wil ik Klaas bedanken, niet alleen voor de samples, maar ook voor het altijd geïnteresseerde en meedenkende oor vanaf de andere kant van de stad! En Michael voor het voldoen aan de verzoeken om telkens weer een emmer sample in de schijnbaar bodemloze IR-put te gooien! Zo ongeveer het hele proefschrift zou zonder jullie bijdragen niet tot stand zijn gekomen.

Ondanks dat de experimenten aan Rhodopsine niet geheel succesvol zijn verlopen, wil ik Arthur en Wim toch bedanken voor hun prettige samenwerking. Het zou zo mooi zijn geweest….

Nothing is more changing as the coffee room… I am not referring to the mice population or the amount of different coffee producing machines that have appeared there. Thanks to everybody for the 'gezelligheid' and hospitality, and hope to see you again somewhere on this globe: Manolis, Mikas, Delmar, Sofia, Natalia, Alessandro, Olga, Magda, Frank, Raoul, Bas, Jante, Anjali, Markus, Krzysztof, Rudi, Maxime, Daniel, Tjaart, Julien, Alisa, Mariangela, Lavanya, Dmitrij, Elena, Denitsa, Chantal, Cecilia, Kostas, Andreas, Silvia, Herbert, Jan and John.

Also special thanks to the reading committee for their attention

and interest they demonstrated while reviewing and commenting the thesis, and all the people who proof read one ore more chapters: Jason, Miguel, Kate, Manolis, Alisa, Toh, Alexander en Thomas.

And thanks to my (ex-) roommates Danielis, Eli, Kinga, and Alexander. It was nice sharing a room with you!

Verder nog een speciaal bedankje voor Miguel: ik weet dat ik altijd op jou kan rekenen. Het ga je goed in Utrecht!

Un gran' abbraccio anche per il mio fellow-italiano: Cosimo. Grazie per la tua amicizia! Spero di venirti a trovare 'laggiú' una volta, cosí mi potrai far vedere come si fa a prendere, pulire, e mangiare una morena!

Back to the international reality of the Biophysics department. I witnessed a big change in the last few years, from a pretty dominant Dutch habitat to an almost completely international one… And also a big change from the boterham met kaas or kroket to all those different dishes that are undoubtebly better then the immensely varied Dutch kitchen… Sterkte Thomas, als laatste lid van het bolwerk tegen de onophoudelijke culinaire en weer-gerelateerde aanvallen!

Vorrei inoltre ringraziare le mia famiglia di Correggio: grazie per il vostro continuo interesse e supporto morale!

Dan wil ik ook mijn familie bedanken, pap en mam en natuurlijk ook mijn zussen Yvonne en Micha: bedankt voor jullie voortdurende stimulans en steun! Zonder jullie allemaal zou ik niet zijn wie ik ben, en ik ben jullie dankbaar voor alles wat ik van jullie geleerd heb, en zal blijven leren!

Da ultimo, vorrei ringraziare la persona piú importante: la Dani! Grazie del tuo supporto continuo, grazie perché ci sei! Questo libro é anche merito tuo!